No. 447
1957

Erwin Roy John

Vol. 71
No. 18

Psychological Monographs

General and Applied

Contributions to the Study of the Problem-Solving Process

By

Erwin Roy John

Department of Physiology, University of California Medical Center, Los Angeles

Price $1.50

Edited by Herbert S. Conrad
Published by The American Psychological Association, Inc.

Psychological Monographs: General and Applied

Combining the Applied Psychology Monographs and the Archives of Psychology with the Psychological Monographs

MANUSCRIPTS should be sent to the Editor.

Because of lack of space, the *Psychological Monographs* can print only the original or advanced contribution of the author. *Background and bibliographic materials must, in general, be totally excluded* or kept to an irreducible minimum. Statistical tables should be used to present only the most important of the statistical data or evidence.

Preparation of manuscripts for publication as monographs should follow the procedures given in the APA *Publication Manual*, 1957 Revision—this revision supersedes all previous manuals for preparing manuscripts for publication as monographs.

CORRESPONDENCE CONCERNING BUSINESS MATTERS (such as subscriptions and sales, change of address, etc.) should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D.C. Address changes must arrive by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee third-class forwarding postage.

Vol. 71, No. 18 Whole No. 447, 1957

Psychological Monographs: General and Applied

Contributions to the Study of the Problem-Solving Process[1]

Erwin Roy John[2]

Department of Physiology, University of California Medical Center, Los Angeles

I. Method

Psychologists have devoted a great deal of attention and effort to the investigation of human problem-solving, because of the theoretical value and practical utility of advances in this area of knowledge. Most of the studies in this area have focused on the success-failure aspect of performance, rather than on the internal structure of the performance. Behavior has been studied in terms of products rather than processes. The time taken to achieve a product (i.e., to solve a problem), or the correctness of the product, or similar static indices have traditionally been the criteria by which performance was evaluated. A severe limitation on our level of understanding of such behavior has been imposed by the use of such methodologies; we are denied access thereby to information about the way in which factors within the performance itself contribute to the outcome, and the way in which these factors are related to other variables than problem-solving performance.

It is possible to achieve solution to a problem efficiently or inefficiently or even fortuitously. It is possible to fail brilliantly. In order to understand the dy-

[1] The author wishes to acknowledge his indebtedness to a number of people who participated in various phases of this work. The Problem-Solving and Information Apparatus (PSI) was devised in collaboration with Dr. Horatio J. A. Rimoldi, now of the Department of Psychology, Loyola University, Chicago. Dr. Rimoldi also participated in the gathering of the data for the Ph.D. population.

The mode of analysis and scoring of PSI performances which is presented was developed with the help of Mr. Martin Balaban, of the Psychology Department of the University of Chicago, and Mrs. Margaret Labadie. The analysis of the data relating PSI performance to various personality measures was performed by Mr. Sidney Blatt, of the Psychology Department of the University of Chicago, who contributed in addition many ideas of value in the course of the research.

I wish to thank Dr. James G. Miller, of the Mental Health Research Institute of the University of Michigan, for enabling this research to be carried out, and for his enthusiastic participation in various phases of its development.

I am indebted to many of my colleagues for constructive criticism: Dr. Benjamin Bloom of the Examiner's Office of the University of Chicago, who gave me access to much useful data about entrance and course examination performance of many of my subjects, as well as to his editorial ability. Dr. Joseph Kamiya, Dr. Martin Deutsch, Dr. Cynthia Deutsch, Dr. Samuel Sutton, Dr. Morris Stein, and Mr. Jo Banks have all aided my attempt to present this material with clarity.

Finally, I wish to acknowledge my debt to my wife, Dr. Vera John, for her assistance in formulating ways to examine the data, for criticism of the manuscript, and, most of all, for making available to me her knowledge of the area of human problem-solving research.

Some of the work included in this paper was done in connection with research supported by the Office of the United States Army Surgeon General under Contracts No. DA-49-007-MD-575 and DA-49-007-MD-684, Dr. James G. Miller, Principal Investigator. The opinions expressed in this publication are not necessarily those of the Office of the Surgeon General, Department of the Army.

[2] Performed this work while at the Committee on Behavioral Sciences, University of Chicago, and Mental Health Research Institute, University of Michigan.

namics of performance, to analyze the factors which relate to the achievement of a solution in a particular way, to investigate the possibility of relationships between variables external to the problem-solving performance and certain aspects of the process, it is desirable to make accessible for direct observation the problem-solving process itself. This requires the development of a technique which will permit the externalization of the dynamics of the process, such as the relationship between the product and the state of information of the subject, instead of continued reliance primarily on techniques which observe product and rely on introspection or subjective report to relate that product to its antecedent causes.

Recognition of this is implicit in some of the theoretical work in the problem-solving area by Duncker (2) and others who have attempted to analyze problem-solving as a process in order to understand the relationships between successful performance and factors within the behavior itself. Theoretical and conceptual schemes have been developed about typical behavior in various situations, and facilitation or hindrance of success. These attempts have been constrained by reliance on recorded verbalizations or qualitative descriptions by an observer of the behavior. In recent years, there have been a number of promising attempts to develop an experimental methodology which would permit more operational definition of the elements of behavior and more precise quantification of these elements. To the extent that such attempts are successful they will enable development of theoretical schemes which can be subjected to experimental verification.

Recognition of the necessity and the desirability of analyzing process rather than product is explicit in the excellent monograph on problem-solving by Bloom (1). The reader is recommended to read this work, if he is not already familiar with it, since it contains not only cogent arguments for preoccupation with devising process measures but a large number of experimental findings which are confirmed in more quantitative terms by the present work.

Some attention has been given to various versions of the game of Twenty Questions, by Harlow (5) and a number of other workers, as a source of process information. These approaches have arisen from concern with military training devices or trouble-shooting analyses, and have tended to deal with procedures for the repair of faulty equipment of one sort or another, where checking procedures are assumed to be binary operations which eliminate each time one-half of the remaining alternatives. While such inquiries are of interest, their utility tends to be limited by the structure of problems which permit such binary operations. Of perhaps more general interest are techniques developed for very similar purposes but without this restriction in structure, such as the Multiple-Alternative Symbolic Trouble Shooting Test (MASTS) of Grings et al. (4), which is yielding some information about previously inaccessible aspects of the trouble-shooting process, in electronic problems. Again, the highly specialized nature of the task limits the utility of results so far obtained for understanding of more generalized problem-solving processes.

Of greater generality, and therefore perhaps of greater potential utility in the investigation of the problem-solving process, is the Tab Item technique, recently developed by Glaser et al. (3). While this task was developed for the analysis of electronic trouble-shooting proficiency, like those above, it possesses a flexibility which permits its application to many other areas of inquiry. The technique consists of presenting a subject with a problem and a set of labeled items of information potentially of utility in the solution of the problem. These items are selected by the subject in any order which seems pertinent to him, until he has gathered those items which he requires for the solution of the problem. The sequence of selection of these items by the subject is used as the index of the way in which already acquired information is used. The authors of the Tab Item indicate their recognition of the flexibility of the technique in areas of more general concern than electronic trouble shooting when they say "Proficiency in medical and psychological diagnosis might be tested by adapting this technique to measure the ability of a clinician to perform the procedures necessary for correct diagnosis. . . . In general, the Tab Item is an applicable technique for the measurement of behavior which involves the serial performance of a set of procedures where the performance of one procedure yields information which supplies a cue for the selection of the next and subsequent procedures." Rephrased, this technique might be of utility in the investigation of problem-solving where the information derived from examination of one aspect of the problem can be used to direct subsequent inquiry to specific other aspects of the problem until those aspects necessary for a stipulation of the solution to the problem are understood.

Working independently, Rimoldi (7) has developed and applied a technique very similar to the Tab Item for the investigation of the process of solution of problems of medical diagnosis and chemical analysis. As in the Tab Item, this technique consists of presenting a problem to a subject and then making available a body of information, containing items which are more or less relevant to the solution of the problem. These items are labeled so that information concerning specific aspects of the problem can be elicited as the subject chooses. The order and nature of the items selected by the subject are recorded.

While the tests referred to above represent a large step forward in problem-solving methodology, they nonetheless appear to the present writer to suffer from a number of drawbacks. As used to date, they are restricted to problem-solving which requires a high degree of specialized knowledge, with the consequence that the process observed is dependent on the education of the testee to an undesirable extent. In addition, it is exceedingly difficult to stipulate the information content of the items available to the subject, or to equate the information content of these items, or to evaluate the extent to which each item contributes to an understanding of the problem as a whole. Perhaps most important, *the format of these problems is not realistic. They tend to structure the path of the subject to an undesirable extent by supplying a limited number of obvious alternatives, give away information by listing the possible alternatives and crucial cues and checks which can be performed, and remove the additional dimensionality which results from having the subject actually deal with real events, rather than their abstract paper representations, in a situation where repetition is possible, and where the set of possible actions is not artificially delimited or ordered in time.*

For greater utility, the ideal format for observation of the problem-solving process should start the subject with a standard minimum of information about a problem and then require him to structure his own presolution behavior with a minimum of externally imposed constraint. Such a format should be based on a task which is maximally free of special skills, special knowledge, or experiences peculiar to a given culture. Since it is important that repeated administration of comparable forms be possible, so that changes in process after various kinds of intervention can be measured, it is also desirable that the effect of familiarity with the generic task be minimal. It is vital to be able to quantify the difficulty of the task, and to be able to vary the difficulty over a large range, as well as to know the information contained in any element of the behavior. Finally, it is desirable to devise a format which presents the subject with the necessity to interact with real events rather than abstract.

A technique which permitted the direct observation of the problem-solving process and met the above stipulations would offer a number of advantages. It would permit analysis of the problem-solving process itself to see whether it was qualitatively homogeneous, or whether it is composed of qualitatively distinct phases which could be defined and described. Theoretical formulations, such as those of Duncker, for instance, could be evaluated against such information and modified if needed. The factors determining the relationship between such phases, and affecting the transition from phase to phase, might emerge from such observation. The relationship between education, experience, cultural background, personality variables, and the internal factors which interact in the problem-solving process might be clarified. The extent to which particular aspects of the process are characteristic of an individual might be determined. The reorganization in process which might result as a consequence of change in the severity of the demands of the task on the individual could be analyzed. The extent to which various techniques or experiences are of value or of utility in increasing the effectiveness of an individual's problem-solving behavior might be evaluated, and knowledge gained about how to change specific aspects of the behavior. Relationships between such aspects of behavior and skills or abilities involved in real-life activities might become evident, enabling better direction of the acquisition of these skills. Diagnostic examination of the problem-solving behavior of an individual might permit application of remedial techniques to the areas of process which most constrain that individual's performance.

The ultimate realization of these possibilities would require extensive research, once suitable instruments for the observation of process were devised. The achievement of the ideal format described earlier represents a difficult task in itself.

This paper describes a technique, briefly reported previously (6), which it is felt constitutes a step in the correct direction. While numerous criticisms can be directed against this technique in that it does not fully meet some of the desiderata, yet it provides much information previously inaccessible and should be of use in the development of still more appropriate methods, and in the acquisition of basic knowledge about problem-solving processes.

In considering the results obtained to date using this technique, reported subsequently in this paper, we must emphasize that we are here dealing with problem-solving performance of a particular kind. *In these problems, a finite and enumerable set of discrete relationships must first be elucidated and then manipulated so as to achieve a single stipulated result.* There clearly are problems which do not fall in the category represented by our technique; problems such as those involved in certain kinds of research in the sciences would appear to be rather similar to our generic task, while problems such as those involved in the making of artistic products, where there is no right answer, or the elucidation of less discrete relationships using probabilistic evaluations, would appear to be quite dissimilar, although perhaps analogous methods can be devised for their investigation. One should not fall into the error of assuming that performance on this task represents evidence of some sort of basic ability which is involved in all sorts of productive activity. This word of caution is made necessary by the many examples of misapplication of "tests" of reasoning ability in selection procedures currently.

Further, there is an unquestionable element of stress and competitiveness inherent in the mode of presentation of the task, as will be described in the following pages. Evidence that subjects are susceptible to this stress and consequently display a performance which by no means represents their optimal behavior has been gathered in a preliminary experiment, and imposes a further constraint on conclusions about the relationship between "real" ability and the sample elicited in our task.

While we will present evidence that behavior in the situation which we have constructed can be interpreted as an approximation of the subject's habitual approach to problems which fall in this category, and is meaningful as such, yet it is apparent that by appropriate instruction or experience over a period of time, or by diminution of the threatening aspects of the situation, this habitual approach can be modified.[3] The extent to which such modification would be generalizable to analogous real-life tasks is a problem of great practical importance.

However, in our desire to qualify the interpretation of results gathered using this method, let us not obscure our belief that we have here defined a real and important class of problem, which would appear to be relevant to a number of

[3] Three individuals were tested in the PSI situation separately on one problem. They were then brought together into a group and asked to solve a second problem. The procedure was for each individual to hand to the experimenter the next step which he would make if he were working alone, and then for the group to discuss its next move until unanimous agreement was achieved. The situation compelled the subjects to formulate verbal arguments for their proposals which would persuade the other members of the group. In this situation radically different behavior was exhibited by the group members from their previous performances, which did not compel the use of language with the consequent imposition of syntactical structure on their "reasoning." This experiment was carried out with the collaboration of Mr. Richard Mann of the Mental Health Research Institute of the University of Michigan.

real-life situations, and that this method gives access to information about problem-solving behavior previously unavailable.

Description of the Problem-Solving and Information Apparatus (PSI)

Problems can be conceived of as consisting of two aspects: elements and relationships. The logic of the relationships between elements can be stated in terms of symbolic logic, devoid of factual content. Thus it is possible to construct an abstract logical problem with the same formal relational structure as a problem which involves factual content. Any logical relations, or propositions, which can be stated in symbolic logic can be stated as an equivalent electromechanical circuit. It is possible, if one confines oneself to problems of a structural complexity commensurate with the number of circuit relationships available, to construct an apparatus on which a large number of sets of logical propositions can be stated. The difficulty (complexity) and structure of these sets of propositions can be quantified as desired, and easily varied by the experimenter. A given set of propositions constitutes a "problem." The achievement of a particular output from a network of electromechanical circuits, which are related in accordance with such a set of logical propositions, requires that the constraints imposed by the logical relationships as defined by that set of propositions be satisfied. The achievement of such an output constitutes a solution to the problem.

In accordance with the foregoing considerations, an apparatus has been constructed (6) and a technique devised which permits observation of aspects of the problem-solving process which fall into three categories:

1. Factors which relate to the amount of effort required and the characteristic habits of work displayed in the process of achieving the desired output from the apparatus.
2. Factors which relate to the acquisition and handling of information in the process.
3. Factors relating to the organization, manipulation, and synthesis of acquired information in order to achieve the desired output.

This apparatus will be designated as the Problem-Solving and Information Apparatus (PSI) in what follows.

Using the PSI, numerous aspects of an individual's problem-solving performance can be defined and measured quantitatively in each of these three categories. The definitions of these variables will be presented in the next section.

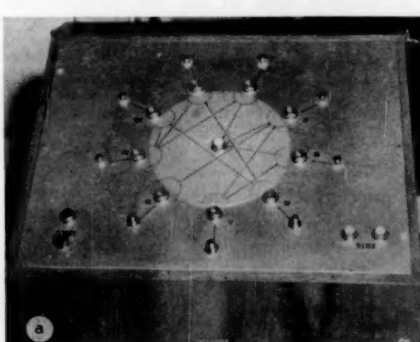
Information theory permits the precise quantification of the total information content, or "difficulty," of any set of propositions, or "problem." One can construct a number of problems of the same or graded difficulty. Such problems can be administered serially to ascertain the effect of increasing difficulty, or the extent to which learning occurs in this situation. The technique permits sequential analysis, making possible a treatment of problem-solving as a branching stochastic process with nonindependent transition probabilities. In other words, one can determine the extent to which the sum of all the information obtained by a subject in the first $n - 1$ steps of a given performance influences the nth step. The utility of this will be discussed further in the section on variables.

The PSI itself consists of a set of electromechanical elements arranged, with other components, into a Boolean computer. Desired logical relationships between the elements can rapidly be established by means of a plugboard in the rear of the apparatus. A given set of relationships defines a problem. Problems are constructed so as to have a unique

solution. A solution is defined as the production of a particular output from the network of related elements by using three particular elements, called "input elements," in some combination or temporal sequence. In other words, the apparatus can be conceived of as composed of three divisions: input elements, an output, and a network relating the input elements to the output. The task of the subject, then, is to learn how to achieve a stipulated output from the network of related elements, by using only the input elements. This requires that he analyze the logical relationships between the elements and then utilize these relationships to produce the desired output. In all of the problems constructed to date, the required output and the set of input elements have been the same. Problems differ only in the structure of the network of relationships between the input and output. Front and rear views of the apparatus are presented in Fig. 1.

On the face of the PSI is a display panel. On this panel is presented a circular array of nine pushbuttons paired with a circular array of nine lights. There are nine basic electromechanical elements in the network, and each light indicates the state of one of the elements. The corresponding pushbutton permits the activation of that element. Inside the circle of lights is a disc on which is presented an array of arrows connecting various elements. An arrow between two elements indicates the existence of a relationship between those two elements, with the direction of the relationship indicated by the head of the arrow. All arrows on the disc stand for relationships in the logical network, and all relationships which exist in the logical network are indicated by arrows on the disc. In the center of the disc is a white light, which is the desired output from the network. Clearly, to each particular problem, defined by a set of relationships, there corresponds a particular problem disc.

Although the existence of all relationship is indicated on the disc, the specific nature of the relationship is not. An arrow from "A" to "B," for instance, might mean (*a*) that A was sufficient to cause B, or (*b*) that A was necessary but not sufficient to cause B, or (*c*) that A was sufficient to prevent B. Sometimes the nature of a relationship can be uniquely inferred simply by inspection of the disc, knowing that all relationships indicated exist and that all relationships which exist are indicated, together with the given fact that all problems are soluble. Sometimes inspection alone does not permit the drawing of a valid inference without ambiguity. In such a case, the subject must devise a procedure to elicit the necessary information from the network. The subject may ascertain the nature of any relationship he desires by activating the pertinent elements by means of their associated pushbuttons and observing the series of consequences of this activation as displayed by the lights on the panel. In other words the subject designs small logical experiments, from the interpretation of which he may infer the nature of the relationships among the elements. *He may use all the elements as many times and in as many combinations as he desires, in order*



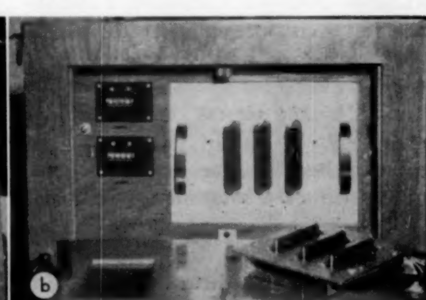


FIG. 1: *a*. Front of PSI apparatus, with example disc in place. *b*. Rear of PSI apparatus, showing plugboard and problem ready to be inserted.

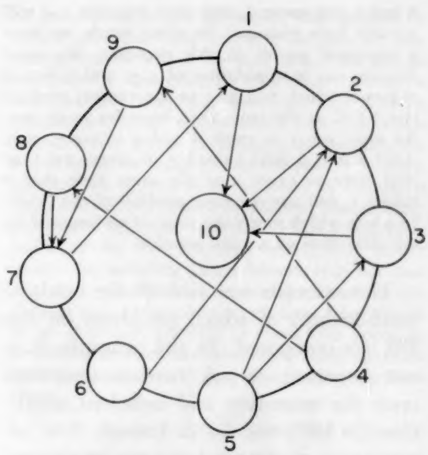


FIG. 2: Example problem disc.

to obtain the information which is necessary to permit the achievement of the required output using only the permitted input elements. Observation of the order, nature, and time of these experiments is the raw data for the sequential analysis referred to above and described in the next section.

Perhaps an example may be useful to clarify the task which the subject faces and to illustrate the way in which information can be elicited from the apparatus.[4] Figure 2 shows the problem disc which has been used to date to accomplish familiarization of the subject with the PSI.

The task of the subject is to learn how to light 10 by some combination or sequence of manipulations which involve using only 4, 5, and 6. In order to learn what this combination or sequence should be, the subject may ask questions using any of the 9 elements in any order or combination. He may also, if he wishes, not ask questions which use elements other than 4, 5, or 6, and attempt to achieve 10 simply by permutations of these three elements. In the procedure which was used in the work reported later in this paper, subjects were asked to achieve solution in minimum time and using minimum questions of the machine, in which case permutation is inefficient, although permitted, behavior.

Let us work the example to illustrate one way in which the subject might proceed. There are three arrows which lead to light 10, the required "output." They come from 1, 2, and 3. These arrows might mean that *1 alone, 2 alone, 3 alone, 1 & 2 but not 3, 1 & 3 but not 2, 2 & 3 but not 1*, or *1 & 2 & 3 together* is the necessary and sufficient condition (NASC) for 10 to occur; in accordance with the permitted relationships as enumerated on page 6, clearly one of the foregoing *must* be true. In order to ascertain which of these possibilities applies in this case, we must ask questions of the PSI. Pressing the pushbutton *1* on the panel will cause the light which indicates the state of *1* to be lit on the display panel. If *1* were the NASC for *10*, after three seconds *1* would go out and light *10* would be lit on the panel. If *1* is not the NASC for *10*, after three seconds light *1* will go out and no light will be lit on the panel. In our example, the NASC for light *10* to be lit is *1 & 2 but not 3*. This means that *1 & 2*, if lit together, will cause *10* to light if *3* was not also lit. By the appropriate manipulations of the PSI, the NASC for *10* are thus determined. From this point on, it is possible to solve the example without the necessity of asking any further questions. Let us see how this can be accomplished, and at the same time let us illustrate how the necessary information for the solution can be achieved by asking questions of the PSI, if one does not draw the inferences already possible.

Let us assume that we have asked the necessary questions to learn that *1 & 2 but not 3* is the NASC. for 10. We now must learn how to light 1 & 2 without 3 by some combination or sequence which involves only 4, 5, and 6, to meet the criterion for a valid solution. Let us first consider how to achieve the ability to light 1 in this way. Only one arrow comes to 1, and it comes from 7. If 1 can be lit by any element, that element must be 7. Since this example, like all permitted problems, is stated to have a valid solution, 7 *must* cause 1. If we do not draw this inference, we can ask what the arrow from 7 to 1 means by pushing button 7. Light 7 will be lit, and three seconds later will go out as light 1 is lit. Three seconds later light 1 will go out, and 10 will not be lit, because 1 is necessary but *not* sufficient for 10. The problem of how to light 1 can now be restated as the problem of how to light 7, since we have learned that 7 is sufficient to cause 1. Two arrows come to 7, one from 8 and one from 9. This might mean that *8 alone, 9 alone*, or *8 & 9 together*, is the NASC for 7. We can, if we wish, ask the PSI which of these is true by pushing the various combinations and observing the results. However, if we look one step further, we see that only one arrow comes to 8 and only one arrow comes to 9. *Both* of these arrows come from 4. If the arrow were to stand for the relationship *prevents*, it would be meaningless because no other element is permitted to

[4] PSI is commercially available, with or without automatic recorder, from Standard Electronics Company, 6728 S. Halsted Avenue, Chicago, Illinois.

cause 8 or 9 within the rules, since no other arrows are indicated. We have stipulated that all arrows on the disc must stand for real relationships. Thus we can rule out the possibility that 4 prevents 8 or 9. Similarly, 4 can not be necessary but not sufficient for either 8 or 9 because there is no other arrow to these elements. Thus, 4 must cause *both* 8 and 9. Now, if it is impossible to achieve 8 without 9, and vice versa, we can rule out the possibility that either one can prevent the other from causing 7, because if that were the case this problem would have no solution. Thus we conclude that *8 & 9* is the NASC for 7, which is the NASC for 1, and we can achieve 8 & 9 from 4.

We have now learned how to achieve one part of the NASC in a fashion which is permitted in the requirement for a solution, which the reader will recall prohibits the use of any other elements than 4, 5, and 6. We must now learn how to achieve 2 without 3, in a similar fashion. Two arrows come to 2, one from 5 and one from 6. This might mean that *6 alone, 5 alone,* or *5 & 6 together* is the NASC for 2. We can ascertain which of these relationships actually holds by asking the necessary questions of the PSI, or we can take an additional factor into account. We know that 3 will prevent 1 & 2 from lighting 10, since the NASC for 10 is *1 & 2 but not 3.* Only one arrow comes to 3, and it comes from 5. If this arrow meant *prevents,* the relationship would be meaningless, since no other element can cause 3. Similarly, it can not mean that 5 and some other element are together necessary for 3 to occur, because there is no other arrow to 3. Thus, 5 must cause 3. If this is true, then if *5 & 6 together* is the NASC for 2, 2 will always be accompanied by 3, and the problem cannot be solved. Since we have stipulated that the problem is soluble, 6 must be the NASC for 2. We can not infer validly whether or not 5 also is sufficient for 2, or whether 5 prevents 6 from causing 2, on the basis of the information we possess. This information is irrelevant, since we can solve the problem without the necessity of resolving this ambiguity. Of course, the ambiguity can be resolved simply by pushing 5 alone and observing the consequence, which is that 2 is lit by 5 as well as by 6.

We now know the meaning of every relationship in the network. We know how to achieve 1 using only the elements permitted in a solution, and similarly we know how to achieve 2 without 3. To achieve 1 we must use 4, which will cause 8 and 9, which will cause 7, which will cause 1. To achieve 2, we must use 6, not 5, because we must prevent 3 from accompanying 2. However, we must achieve 1 and 2 at the same time, since 1 and 2 together is the NASC for 10. A moment of thought will make clear to the reader that if 4 and 6 are pushed together, after three seconds 8 and 9 and 2 will be lit, then 2 will go out and 8 and 9 will cause 7, and when 7 causes 1, 2 will already have gone out. In other words, we have a temporal aspect to this problem. We must arrange our manipulation of 4, 5, and 6 into a sequence which will give us the various parts of the NASC *at the same time.* In order to do this, we must use 4 to cause 8 and 9 to occur, wait until 8 and 9 have caused 7 to occur, use 6 *at that time* to cause 2 at the same time that 7 causes 1, and now we have established the NASC in a way which meets the restrictions imposed by the definition of a valid solution.

The example contains all the kinds of relationships of which problems on the PSI are composed. In the example, it is not necessary to ask further questions once the necessary and sufficient condition (NASC) for 10 is known. Not all problems, of course, have this characteristic. An enormous number of different problems can be constructed from the various meaningful combinations of these relationships which the design of the PSI permits. The information content of these problems can be rigorously evaluated, and the content of each possible question which a subject might ask of the PSI can also be evaluated, in a fashion which will be discussed in detail in the following pages.

It should now be apparent that the subject, as a result of his manipulations of the pushbuttons in the course of investigating relationships, is presented with ordered sequences of events which are consequences of his activation of elements. His task is to extract, from observation of the sequences which he designs, that information which he needs in order to be able to ascertain the characteristics of a valid solution to the problem. Such a solution may require an exceedingly complex ordering and combination of the input elements.

Note: The three-second delay which ensues before the consequences of any antecedent are displayed has been imposed to give the subject the opportunity to see the full set of consequences of any change of state which he initiates,

as an ordered series of events, and as well to afford sufficient time for him to execute any manipulations which he may desire to add at any point in the series. In preliminary studies, a seven-second delay was provided and subjects complained about the excessive delay. A three-second delay seems empirically to provide the opportunity to order events adequately without causing the subject to be annoyed by having to wait, and is considered a satisfactory compromise. Because of the large variation in rate of manipulation of the PSI, described under *Rate* in the section on variables which follows, it is felt that the three-second delay is not the cause of the various rate phenomena which have been found. It is, of course, possible to check this point by providing a manually controlled delay operated by the subject at will, but this has not yet been done.

Some readers may be familiar with the models of neural nets which have been extensively used by neurophysiologists and cyberneticists. It may be of assistance to conceptualize the PSI as an electromechanical neural-net model, in which direct excitation, summation, and inhibition can be combined in any desired fashion. Or, perhaps, the reader may appreciate further clarification from Fig. 3, which is a verbal equivalent of a problem on the PSI, as it could be represented once the subject has elicited all the information necessary to stipulate each relationship. Note that once the subject has obtained knowledge about all the relationships he is still left with the nontrivial task of utilizing this knowledge to produce a solution of the problem as required by the rules.

Familiarization with the PSI is achieved by stating the rules, describing the nature of the possible logical relationships which the subject might expect to encounter later, and then illustrating all of these by presenting the comprehensive detailed example which we have just completed. The subject is provided with paper and pencil, and encouraged to take notes. Young children are able to work effectively on the PSI, indicating that the situation itself offers little difficulty of comprehension when illustrated concretely instead of in the manner used above.

PROBLEM 1

IF IT IS TRUE THAT:

1 alone causes nothing — 6 alone causes nothing

2 alone causes nothing — 8 alone causes nothing

3 alone causes nothing — 9 alone causes nothing

4 alone causes 7 plus 8

5 alone causes 2

7 alone causes 3

3 plus 5 causes 1 plus 2

6 plus 8 causes 9

5 plus 9 causes nothing

3 plus 5 plus 9 causes 1

1 plus 2 causes nothing

2 plus 6 causes nothing

1 plus 6 causes 10

1 plus 2 plus 6 causes nothing

THEN,

CAUSE 10, BY SOME COMBINATION OR SEQUENCE WHICH USES ONLY 4, 5, and 6

FIG. 3: Verbal equivalent of first problem. Note: This represents the problem once all relationships have been analyzed.

II. Procedure and Scoring

A. General Considerations

A large amount of data is made available from a single performance on this apparatus. There are many possible ways of organizing and examining these data. As will be apparent from the subsequent list of variables, it is not difficult to define a variable which is a new way of looking at some aspect of the data. The possibilities for such new variables have by no means already been exhausted. Particularly, the definition of variables which give access to the dynamic aspects of our data without emphasizing only the static aspects has been difficult. Our habitual approach has been in terms of

products rather than process, and, presented with precise and detailed data about process, one tends still to describe the product rather than the process. It is hoped that, as work in this direction proceeds, a set of more adequate process indices will be developed.

Many of the variables defined are reasonably independent, many covary appreciably with others. In view of the differences in performance of groups of diverse training and background, which will be illustrated in the data to be presented, it would appear unprofitable to investigate the intercorrelations between these various variables at this time. The evidence suggests strongly that one should use a homogeneous population in order to determine intercorrelations between the variables, in view of the probability that the internal organization of the process, and therefore the correlations to be achieved, will vary as the structure of the group varies. Changes in situation and instruction may also change the internal organization. This pilot study did not attempt to obtain such information. Perhaps ultimately factor-analytic methods should be used to select a set of variables which span the space of the performance. However, the intuitive meanings and significance of many of these variables have been so useful conceptually that we have made no effort to prune our set of variables to a minimum. Such an attempt would appear premature until an appreciable period of examination of the many different methods of organizing this somewhat unfamiliar kind of data has elapsed.

The variables so far defined and examined can be grouped into three different major aspects of performance:

1. Those variables which describe the individual's work habits and relate to the effort expended in the solution of the problem are referred to as *work variables.*

2. Those variables which describe the way in which the individual acquires and handles information are referred to as *information variables.*

3. Those variables which describe the orientation of the individual to the problem, and which relate the effect of acquired information to this approach, are referred to as *approach variables.*

B. Problems

Subjects were presented with two problems on the PSI. Both problems were administered to each subject separately in a single session, which usually lasted about one and one-half hours, including time for familiarization and demonstration of the example.

The two problems were constructed so as to be almost identical, with one major difference distinguishing them. Each contained 17 items of information in the total information pool; that is, 17 simple propositions would define the network of relationships exhaustively. Each problem contains 4 relationships which are direct implication, 2 relationships which are conjunction, 1 relationship which is compound conjunction and disjunction (x will occur if and only if y and z but not w has occurred), and 1 relationship which is disjunction. The difference between the two problems is that the first problem requires that *three* coincidences be achieved in the course of the solution, while the second problem requires that *four* coincidences be achieved in the course of solution. One of the four coincidences used in the second problem is the same relationship which is required to produce a prior coincidence, but this relationship is used for a different purpose in the two instances. Thus the difference between the two problems might be summarized by stating that they are identical except that a relationship used only once in the

first problem is used in two different ways in the second problem, as a consequence of using different elements in the same relations.

Attention is directed to the apparent small difference between these two problems, since presumably the differences in performance observed are a consequence of these minimal differences in structure. Yet *this minimal difference in structure,* achieved without any increase in the number or nature of the items in the pool, *brings about a major change in the behavior elicited from our subjects.* The data presented below give us an insight into the extent to which small changes in structure may change aspects of response which superficially one would expect to remain relatively unaffected. Clearly, the *structure* of relationships, as well as the sum of the information contained in these relationships, is a major determinant of the ease with which *meaning,* or coherent organization of the set into a whole which is unified, can be achieved. (One might suppose, since alternative structures may lend themselves to the same content, that clarification of this point might generate understanding of the factors which constrain comprehension under various conditions of the presentation of information.)

C. Instructions

In the experiments which yielded the data to be here reported, subjects were asked to solve the problem in a *minimum* of time and with a *minimum* number of inquiries of the PSI. The evidence strongly suggests that other instruction, for instance to minimize time alone or inquiries alone, or other situations, which increased the probability that the subject would use particular mediational processes such as language, might lead to a different approach to the situation. This particular choice of instruction was made in an attempt to place the subject under some pressure to perform effectively while at the same time minimizing the extent to which the situation was structured. It was hoped that this would enable us to observe the way in which our subjects themselves tended to structure cognitive tasks of a highly abstract nature.

It should be clear that the PSI permits the construction of other situations, and one can then compare the behavior elicited in these situations to that obtained under the conditions cited above. Specific content can be put back into the problem, instead of using an abstract presentation. Problems can be presented with all relationships indicated and their nature specified. Problems can be presented with none of the relationships indicated. Networks can be constructed which are not constant, but which change as a function of certain aspects of an individual's performance. Other instructions can be used. Betting behavior, or behavior under stress, or under various kinds of motivation can be observed. Group situations can be studied. Much can be learned about the dynamics of problem-solving behavior, and other kinds of behavior as well, by such variations. None of these interesting alternative approaches has been studied to date. In view of the many possible variations in the situation in which the PSI is administered and in the amount of content introduced into the particular problems selected, the information which has been gathered to date must be regarded as surveying only a small part of the domain which can be investigated using this basic technique. We will develop certain ideas and approaches to the analysis of this kind of sequentially observed behavior which may be of utility in the direction of subsequent studies. However, the results obtained with regard to differences between the performance of various kinds of groups and with regard to characteristics of the problem-solving process must be regarded as gross outlines which must yet be filled in, rather than as definitive descriptions. This is particularly true in view of the small size of our various groups.

D. Subjects

The subjects in these experiments were 59 University of Chicago students and staff members who responded to an advertisement offering to pay persons who

would participate in a psychological experiment. These students varied appreciably in their length of attendance at the University and came from various academic disciplines. On the basis of the criteria stated below, they were divided into groups as indicated.

Group 1: 21 Ph.D. candidates or recent Ph.D. recipients. (Composed of Group 3 and Group 4, below.)

Group 2: 16 students in their first quarter of residence in the College of the University of Chicago.

Group 3: 10 candidates for or recent recipients of the Ph.D. degree in the natural sciences.

Group 4: 11 candidates for or recent recipients of the Ph.D. degree in other fields (8 social sciences, 3 humanities).

Group 5: 11 first-, second-, and third-year students in the College of the University of Chicago who indicated in an essay on career plans their intention of studying for advanced degrees in the natural sciences.

Group 6: 11 first-, second-, and third-year students in the College of the University of Chicago who indicated in an essay on career plans their intention of studying for advanced degrees in areas other than the natural sciences.

The mean performance of each of these groups will be presented for each PSI variable on both of the problems described in Section B. Our purpose in the remainder of this chapter is to present the definitions of the variables which constitute our present mode of analysis of PSI data, and to gain some insight into the extent to which the desiderata set forth on page 3 are achieved in the PSI. By discussing the relative performance of the six groups of subjects we hope to gain some idea of the dependence of PSI performance on the level of difficulty of the problem, on familiarity with PSI problems, on the level of education of the subject, on the specialized skills of the subject as indicated by technical training, and on the interests of the subject as indicated by career plans.

In a few cases we will indicate correlations with other variables. These correlations will refer only to problem 2, unless otherwise stated. The Examiner's Office of the University of Chicago has permitted us access to the results of various tests administered to students by that office. Among the tests administered were the American Council on Education Psychological Examination and a mathematical aptitude test. In addition, a number of tests were administered which have subtests designated as "analytic" scales. An average score for the combined analytic sections of these tests was computed and will be referred to as Total Analysis. Some of the correlations between these tests and PSI variables will be presented in what follows.

Our concern here, then, is to understand the set of PSI variables, rather than the differences between our six groups.

E. Work Variables

1. TIME—the number of minutes required for solution to be achieved. Solution criterion is the ability to light the center light using only buttons 4, 5, and 6 in some combination or sequence, *on three successive attempts.* (Data for this variable are presented in Table 1.)

Correlations. Time correlates with Effort (see below) .88, with the ACE "T" scale −.36, with the ACE "Q" scale −.37, and with Total Analysis as described in the preceding section −.40.

Discussion. Time is a power index rather than a process index. Certain relationships between Time and process are, of course, to be expected, since the process must occupy time. Inspection of the data shows that the various groups do not differ markedly on the time which they require for solution of problem 1. The difference is appreciably greater on problem 2. Note that groups 2, 5, and 6, which are the college student groups, require less time than groups 1, 3, and 4, which are the Ph.D. groups. As will be seen later, the faster achievement of solution by these

TABLE 1

TIME

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. college)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	23.8	22.0	18.4	22.0	23.0	20.7	21.4
Problem 2	44.5	41.0	31.0	39.2	41.8	37.2	50.0
Difference	−20.7	−19.0	−12.6	−17.2	−18.8	−16.5	−28.6

groups is accompanied by what might be termed lesser economy of inquiry. When the difficulty of the problem is increased, the same differential is manifested with the exception of group 6, the non-natural-scientist college group, which encounters real difficulty with various aspects of this problem. The increase in required time as the difficulty is increased is approximately the same for all groups except group 6. Thus it appears that the effect of education as such on the time required for solution of a PSI problem is to slow the process. The effect of special technical interest or training, as seen in the differential between groups 3 and 4 and groups 5 and 6, appears to cause slightly faster performance (where "technical" is used to denote natural science skills). None of these differences seems particularly large. The over-all correlation between performance on the two problems for this variable is .46. This correlation is not really a reliability index since the two problems are not equivalent in their structure. The increased familiarity of the subjects with the PSI after solving problem 1 does not bring about faster performance on problem 2. An accurate measure of the interaction would require varying the order of administration, which has not yet been done.

2. QUESTIONS—the number of questions required for solution to be achieved.

A question is defined as the series of manipulations of the PSI which occurs between the activation of *any* element in the net and the first subsequent time period in which *all* elements are inactive. In other words, a question is considered to be a manipulation or series of manipulations which interact or are contemporaneous. Both the specific content of each question and the time at which it was asked are recorded. (Data for this variable are presented in Table 2.)

Correlations. Questions (Q) is related to Effort (see following) .97, to Rate (see following) .49, to Actual Redundancy (see following) on problem 1, .72 and on problem 2, .77, to the ACE "T" scale — .43, to the ACE "Q" scale — .35, and to Total Analysis —.46.

Discussion. The number of questions required for the solution to be achieved is also a power index. The nature of the questions, that is, the content of the questions, is a process index. The treatment of this data will be presented in the next section, on Information Variables. The correlation between Q and Effort will be easily understood when the reader realizes that Effort is approximately ½ Q times T. The correlation between Q and Actual Redundancy is of great importance and will be discussed in detail in the presentation of that variable. As with time, the data suggest that there is a reasonably high relationship between the power aspects of PSI performance and ACE "T" and "Q" scales and that the relationship of the achievement of a product on the PSI to the ability involved in the "Analy-

TABLE 2

QUESTIONS

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other) College)
Problem 1	69.1	35.0	51.0	23.0	46.0	57.8	77.2
Problem 2	110.0	50.0	77.0	41.0	59.0	82.3	156.0
Difference	−40.9	−15.0	−26.0	−18.0	−13.0	−24.5	−78.8

TABLE 3
COMPLEXITY

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	1.71	1.67	1.80	1.53	1.76	1.78	1.72
Problem 2	2.30	2.08	2.70	1.98	2.15	2.03	2.55
Difference	−.59	−.41	−.90	−.45	−.39	−.25	−.83

sis" subtests mentioned above is definite, being significant at the .01 level. The correlation for all subjects between performance on problems 1 and 2 is .46, the same as for Time. Note that groups 1, 3, and 4 require fewer questions than groups 2, 5, and 6 on both problems. This indicates that the more educated subjects ask fewer questions than the less educated ones. Note also that group 3 requires fewer Q than group 4, and group 5 requires fewer Q than group 6. Thus it appears that PSI performance using Q as an index is more economical as the educational level of the subject goes up, and, if the educational level is held constant, training or interest in the natural sciences is related to more economical performance. Finally, if we look at the increment in Q which accompanies an increment in the difficulty of the problem, we see that groups 1, 3, and 4 are affected less by this change than the other groups, although group 6 is much more troubled by this increase in difficulty with respect to this index than any other group. This index then is definitely not independent of education, or specialized training or interests. Furthermore, the increase in familiarity from the first to the second problem is not enough to offset the apparent slight increase in complexity.

3. COMPLEXITY—the total number of manipulations required for solution to be achieved, divided by the total number of questions (Q). This is the average complexity over the process as a whole. Complexity is also computed question by question (discussed under Analytic-Synthetic Shift below). (Data for this variable are presented in Table 3.)

Correlations. None computed.

Discussion. Complexity is an index which has more of the aspects of a process measure than the two indices so far discussed. Characteristically, during a performance questions require longer and longer series of manipulations, but are generally single manipulations at the beginning of the performance. The differential complexity of different phases of the process is used as an index of the mode of approach as the process evolves (see section on Approach Variables following). Here we are dealing with overall complexity, which is an average value retaining little of the information inherent in the measure itself as a process index. Examining the data we see that on the simpler problem there is little variation in the mean complexity from group to group. As the difficulty of the problem increases, however, we see that the groups differentiate, with groups 2 and 6 showing an appreciably greater increase in complexity than the other groups. The differential between groups 1 and 2 and groups 5 and 6 suggests that the greater the educational level the less the complexity (1 vs. 2), and the more familiarity or interest in the natural sciences the less the complexity (5 vs. 6). The extent to which the factors of educational level and specialized training or interest affect this index appears to depend on the severity of the demands made on the individual by the problem. Increased familiarity with the PSI does not cause a decrease in the complexity index in this case.

4. RATE—the number of questions required for solution divided by the Time. (Data for this variable are presented in Table 4).

Correlations. Rate correlates with Effort (see following) .37, with Actual Redundancy .57, with Questions .49, with Pauses (see following) −.80, with Mixture of Modes (see following) .08, with Predominant Mode (see following) −.18.

Discussion. Rate is the first PSI variable we have discussed so far which is a process variable. Examination of the data for this index shows clearly that there is an underlying integrity and unity to the process of problem-solving which culminates in the achievement of a product (correlation coefficient between problem 1 and problem 2 is $r = .84$), yet varies greatly from individual to individual.

TABLE 4

RATE

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	3.01	2.02	3.42	1.77	2.25	2.62	4.03
Problem 2	2.27	1.47	2.62	1.17	1.76	1.94	3.10
Difference	.74	.55	.80	.60	.49	.68	.93

If one plots a graph of cumulative questions versus time, which we will subsequently term the Output Graph, the curve so generated is characteristically a straight line, the slope of which approximates Rate. The regularity with which such a straight line is observed suggests that in the PSI situation an individual evaluates the information which he possesses in such a way as *to come to a decision in a characteristic time.* While on the whole the performances can be approximated by a straight line, deviations from this straightness occur in two ways. Occasionally one observes inflection points at which a change in slope occurs (see Inflection Points following), and occasionally one observes pauses in output followed by a transient period of accelerated output until the curve returns to the basic straight line, after which the previous slope is resumed (see Pauses following). In Fig. 4, a number of Output Graphs are presented.

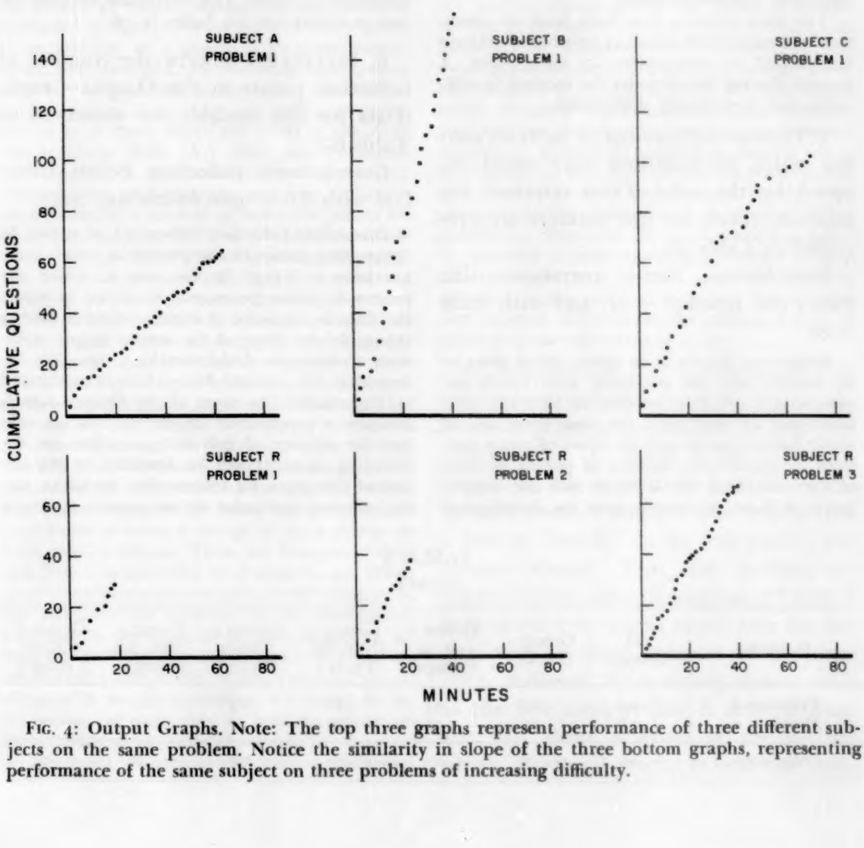


FIG. 4: Output Graphs. Note: The top three graphs represent performance of three different subjects on the same problem. Notice the similarity in slope of the three bottom graphs, representing performance of the same subject on three problems of increasing difficulty.

The three top graphs in Fig. 4 represent three different persons solving the same problem. The top left curve represents an average performance; while it is approximately a straight line, there are a number of minor deviations from linearity. The top middle graph represents a highly linear performance, obtained fairly often (see Changes of Set following). The top right graph represents an unusual performance with multiple inflection points. The bottom three graphs represent the performance of the same individual on three problems of increasing difficulty. Note the similarity in slope despite the increase in difficulty.

From the correlations above we see that as Rate increases, the economy of Effort and Questions increases, the inefficiency of handling information (Actual Redundancy) increases and the frequency of Pauses decreases. The relation of Rate to the two Mode indices shows that speed of asking questions is essentially independent of whether the subject is in the Analytic or Synthetic phase of the process (see discussion under Approach Variables below).

The data indicate that both level of education and specialized technical training or interest are related to performance at lower rates. A general slowing accompanies the increase in difficulty, but is markedly differential.

5. Pauses—the number of minutes during which no questions were asked, divided by the total Time required for solution. (Data for this variable are presented in Table 5.)

Correlations. Pauses correlates with Effort (see below), −.22, and with Rate −.80.

Discussion. Pauses is an index which gives us an insight into the constancy with which the process goes on. It is possible to have two performances which display the same Rate, one of which has no pauses and the other of which contains an appreciable number of pauses, because of the smoothing which enters into the computation of Rate. In the two cases the distribution of actions differs appreciably with respect to grouping. As can be seen from the correlation of −.80 between Pauses and Rate, the relationship between these two is high, yet the evenness of spacing of questions is not the determinant of Rate. Note the large differentials in the above data. We see that educational level seems to affect the frequency of pauses, by contrasting the data for groups 1, 3, and 4 with those for groups 2, 5, and 6. By comparing group 3 with group 4, and group 5 with group 6, we see that at the same educational level the group with "technical" training or interest displays a higher incidence of pauses. This holds true at both levels of difficulty. There is a general increase in the frequency of pauses as the problem increases in difficulty. Note that the higher the initial frequency of pauses, the greater the increase in frequency which accompanies greater difficulty. Thus this index appears to be somewhat dependent on both the educational level and the special skills of subjects. Increased familiarity with the PSI does not result in a decrease in frequency of pauses. The correlation between the two problems for this index is .76.

6. Inflection Points—the number of inflection points in the Output Graph. (Data for this variable are presented in Table 6.)

Correlations. Inflection Points correlate with Effort (see following) .42.

Discussion. Inflection Points are of utility in designating points in the process at which there has been a change in the rate at which the process is being generated. Since we interpret the Rate as an index of decision-making time, a change in the slope of the output graph represents a change in decision-making time. We believe that the constant Rate which is manifested by the straight line form of the Output Graph indicates a *constant set* on the part of the subject. In support of this interpretation are the following facts: As will be described in the section of this paper on Information Variables, one can stipulate the point in the process at which

TABLE 5

Pauses

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	.09	.20	.05	.26	.15	.12	.03
Problem 2	.18	.30	.11	.39	.23	.20	.06
Difference	−.09	−.10	−.06	−.13	−.08	−.08	−.03

TABLE 6

INFLECTION POINTS

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci.) College)	Group 6 (Other College)
Problem 1	1.5	1.1	.9	.6	1.6	1.2	1.4
Problem 2	2.5	1.4	2.2	1.1	1.7	2.2	3.3
Difference	−1.0	−.3	−1.3	−.5	−.1	−1.0	−1.9

the necessary and sufficient information (NASI) for the solution has been achieved by the subject. When inflection points occur in the Output Graph, they are most frequently associated with the achievement of the NASI. That is, when the subject achieves the information which is both necessary and sufficient to permit him to infer the valid solution to the PSI, characteristically there will be a change in the Rate, even though solution may not be forthcoming for some time.

Some figures here may give an indication of the extent to which changes in Rate can be attributed to *changes in the state of information* of the subject. In a group of 90 performances which were analyzed, the following observations were made. Seventeen performances had no inflection points. Forty had one inflection point, and in 33 of these either the NASI or the Analytic-Synthetic Shift (A-S Shift, see following) was located at the inflection point (82.5%). Twenty-three performances had two inflection points, making a total of 46 inflection points for this group, and the NASI or A-S shift was coincident with 25 of the inflection points, or 54.3%. Ten had three or more inflection points, making a total of 40 inflection points for this group, and the NASI or A-S shift was coincident with 15 of the inflection points, or 37.5%. As the difficulty of problems increases, the percentage of such coincidences goes up. Finally, on 27 of the 90 performances, a pause occurred immediately before the A-S shift. These data appear to support the conclusion that the stable Rate observed in the PSI is an indication, and perhaps a consequence, of the existence of a constant set. When information input to the individual occurs which is probable to cause a change in set, a change in Rate usually follows. Thus, we interpret Inflection Points as an index of changes in set. When such changes occur in conjunction with changes in the adequacy for solution of the information possessed by the subject, we take this as an indication that the new information is being incorporated conceptually. When inflection points occur with no corresponding increment in information, we take this to indicate either reorganization of the body of information acquired to this point, or misinterpretation of the meaning of the most recently acquired items. Subsequent steps in the process indicate which of the above is correct, according to their content. We score the appropriateness of "changes of set" by relating the loci of inflection points to the loci of NASI points, in a yes-no dichotomy, as a subscore under this heading.

Note that the data show an appreciable differential between the groups for this variable. While this differential is clearer on problem 2 than on problem 1, it indicates in general that the level of education and the extent of special technical training or interest are both related to frequency of inflection points. Comparing the more highly educated groups (1, 3, and 4) to the less educated (2, 5, and 6) shows the first differential. Holding the educational level constant and looking at the effect of special technical training or interest by comparing natural scientists (group 3) with non-natural scientists (group 4) at the Ph.D. level and similarly (group 5 vs. group 6) at the college level, shows the second differential. Note also the much smaller increase in incidence of inflection points among the more educated groups (1, 3, and 4) as the level of difficulty of the problem increases. The correlation between performance on problem 1 and on problem 2 for this variable is .51.

7. PERCENTAGE OF NONLINEARITY—a measure of the extent to which an inflection point changes the subsequent performance. If there is no inflection point, % NL is zero. If there is an inflection point, a straight line is drawn as a line of "best-fit" to the first part of the Output Graph. This line is then extended to the time of solution. (There is sometimes a transient rapid rate for the first two or three minutes of performance, followed by a steady slower rate. If the first straight line is derived from the first three minutes of the process alone, disregard it and extend the slope

TABLE 7

Percentage of Nonlinearity

Problem	All Groups
Problem 1	12.6%
Problem 2	22.5%

of the performance after the third minute.) The area of the large triangle under this straight line is determined, as is the area between the extended straight line and the actual performance polygon, indicated as the small triangle labeled "deviation from linearity" in Fig. 5. The ratio of the latter to the former is percentage of Nonlinearity. (Data for this variable are presented in Table 7.)

Correlations. Percentage of Nonlinearity correlates with Actual Redundancy (see following) .24.

Discussion. This index is used in conjunction with the previous one in order to estimate the extent to which an inflection point affects the subsequent process. We argue that if a change in set is major, the characteristic decision-making time which appears to be related to set or conceptual framework (see previous discussion) should change appreciably. This index measures the accumulated amount of change in decision-making time summed over the entire process from the point of change and stated as a percentage. We assume that if there were no change in slope of the output graph, the first straight line, extended to the time of solution, would closely approximate the actual performance. To the extent that this performance proceeds at a rate different from this initial rate, the area between the performance polygon and the extended initial slope will increase. By division by the total area under the extended initial slope, we state this deviation as a percentage, the sign of which indicates whether the change in rate was an increase or decrease.

Percentage of Nonlinearity tends to be larger on the more difficult problem. Also, as the % NL goes up, the performance tends to become somewhat more inefficient insofar as utilization of information is concerned. The correlation between performance on this variable from problem 1 to problem 2 is .02. Further investigation with parallel forms of PSI problems, equated for difficulty, will be necessary before one can determine whether the low correlation indicates unreliability of this index, or results from the generally less efficient performance which subjects display on the more difficult problem.

8. Effort—the area under the performance polygon on the Output Graph.[5] (Data for this variable are presented in Table 8.)

Correlations. Effort correlates with Pauses −.22, with Questions .97, with Rate .37, with Actual Redundancy (see following) .72, with Time .88, with Inflection Points .42, with Changes of Approach (see following) .59, with Relative Analytic-Synthetic Shift −.59, with Absolute Inferential Lag .77, with Relative Inferential Lag .30, with ACE "T" scale −.45, with ACE "Q" scale −.33, with ACE "L" scale −.36, and with Total Analysis −.51.

Discussion. Effort was adopted as an index to attempt to incorporate both Time and Questions in one measure. As can be seen from the above correlations, these two measures are sufficiently highly related so that Effort does not tell us much that either of the simpler measures could

[5] Effort was measured by plotting the Output Graphs on standard axes and measuring the pertinent areas by use of an Ott Planimeter.

TABLE 8

Effort

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	379.	155.	299.	67.	235.	298.	385.
Problem 2	815.	408.	462.	321.	488.	568.	1095.
Difference	−436.	−253.	−163.	−254.	−253.	−270.	−710.

not do as well. Yet, because it did combine the two into a single index, a number of correlations were computed using this index. They can be summarized as follows: As effort increases, pauses tend to decrease, rate tends to increase, changes of set or inflection points become more frequent, shift from analysis to synthesis and back to analysis occurs more often, inferences are less validly drawn and tested, and the shift from analysis to synthesis tends to occur relatively earlier in the process. The correlation between performance on the two problems for this measure is .42.

Examination of the data above indicates that this measure is dependent both on educational level and special technical training, as must be expected to be consistent with the dependence of the two components of this index on the same factors. Note the surprising uniformity of the increment in Effort with increased difficulty of the problem for the various groups except group 6.

Note that both ACE "Q" and "L" scales correlate significantly at the .05 level with Effort, which indicates that verbal as well as quantitative skills are useful in PSI performance, yet clearly neither is sufficient to be the major determinant of performance quality. Finally, the high correlation, significant at the .01 level, with Total Analysis indicates that the skills involved in PSI performance are similar to the skills involved in the elucidation and application of principles in other areas more relevant to real-life problems.

F. Information Variables

At each step in the problem-solving process, as we observe it in this special situation, two kinds of information are obtained by the subject:

1. Information overtly explicit in the answer to the questions asked of the machine, referred to as Q_e.
2. Information obtainable by logical inference from the answer to the question asked, taking into consideration the implications of all previously acquired knowledge. This will be referred to as Q_i.

We can represent the process of acquisition of information from the apparatus by an individual as the sum yield of a series of questions, each of which contains in its answer an explicit (Q_e) and an implicit (Q_i) amount of information. It is clear that under certain conditions Q_i will be a function not only of the question just answered, but of information previously obtained. Thus, *the total information content of the answer to a particular question may not be an invariant characteristic of the question, but a function of the sequence of questions asked previously.* The process then can be represented as follows:

Total information available after the *n*th question =

$$\Sigma(Q_e + Q_i)_1 + (Q_e + Q_i)_2 + \cdots +$$

$$(Q_e + Q_i)_{n-1} + (Q_e + Q_i)_n = \sum_{j=1}^{n}(Q_e + Q_i)_j =$$

the sum of all explicit and inferred information.

If the *explicit* information content of the answer to any question, $(Q_e)_n$, is contained in the sum of the explicit and implicit information content from the first answer to the $n-$*first* answer elicited from the apparatus in the process, we will term that question *inferrable.* Obtaining the answer to that question will not further increase the sum of elicited information. It is possible, by changing the ordering of a given set of questions, to vary greatly the number of inferrable questions in the set. If we define the *actual redundancy* as the number of inferrable questions in a particularly ordered set divided by the total number of questions in that set, *it is possible to show that the actual redundancy of a particularly ordered set can be appreciably less than the actual redundancy of that set differently ordered.* Similarly, it can be shown that *the redundancy of a particularly ordered set can be less than the redundancy of another set which contains fewer questions.*

We thus reach the conclusion that if we wish to quantify the efficiency of acquisition of information in a given problem-solving process, it is not adequate merely to consider the number and

choice of questions asked in the process, treating the information content of the answer as only a function of the particular question. In order to state the efficiency of a given process, that is, in order to state the extent to which the information which was acquired up to a given step was used to determine the decision as to what should be the next step in the process, we must recognize that the rate of acquisition of information may well depend on the *ordering* of questions in addition to the *nature* of the questions.

Further, we reach the conclusion that there may be a fortuitous factor involved to some extent in the problem-solving process. Although there may be no a priori basis for asking a particular question first in a series, yet the selection of that question as a starting point, or the asking of that question before some other question but after others, will affect the extent to which portions of the information content of the total pool can be inferred at any point in the subsequent process. (Note: This would appear to have interesting consequences in a number of areas. It would seem that to score an examination or a psychological test, in which the contribution of this kind of factor to the performance has *not* been ruled out, on the basis of the amount of time taken or any similar nonsequential "power" criterion for quality of performance, is to include the possibility that a subject with a poor score is actually more efficient than a subject with a "good" score. Consideration of the residual variance in actual redundancy other than that due to the number of questions [discussed below] shows that *ordering is as important as what is asked*.)

In order to partial out the contribution of this fortuitous factor of initial choice, and in order to enable the precise evaluation of redundancy by accounting for the dependence of information content on ordering, we have analyzed each problem used in this study in the following way: A flow sheet has been developed which shows, *for any possible series of questions,* what the explicit and implicit information obtained at each step will be. The necessary and sufficient information for the solution of each problem and the total content of the information pool for each problem is indicated on each flow sheet. Detailed exposition of such a flow sheet for a typical problem requires too much space to be here presented. Suffice it to say that when such a flow sheet has been constructed, a typical problem-solving performance can be sequentially analyzed for the efficiency of information handling in a few minutes, with absolute reproducibility and precision.

The analysis also indicates the point in the performance where the necessary and sufficient items of information for the solution of the problems have been *explicitly* achieved, subsequently referred to as the *NASI(exp)*, and the point where these could be implicitly achieved if maximum inference were validly drawn, subsequently referred to as the *NASI (imp)*.

Other measures of redundancy can be defined. There are a number of different kinds of redundancy, for example, the number of times in a particular performance that the *same* question is repeated. We have selected this particular index as best suited to our present purpose, but recognize that redundancy can meaningfully be defined in a number of alternative ways, and call to the attention of the reader such potentially useful measures of redundancy as repetition of particular questions or patterns of ques-

TABLE 9
EXHAUSTIVENESS OF INQUIRY

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	13.6	12.4	13.5	11.8	12.9	13.3	13.9
Problem 2	13.0	12.8	13.4	12.1	13.4	13.5	12.4
Difference	.6	−.4	.1	−.3	−.5	−.2	1.5

tions, which are frequently encountered.

We are now in a position to define a number of different variables related to the acquisition and handling of information:

1. EXHAUSTIVENESS OF INQUIRY—the number of different items actually elicited from the information pool. There are, in each of the two problems herein discussed, only 17 possible different meaningful items. All questions can be decomposed into aggregates of these basic items. (Data for this variable are presented in Table 9.)

Correlations. Exhaustiveness of Inquiry correlates with Actual Redundancy (see following) .10.

Discussion. This variable tells us the extent to which the set of possible "basic" questions has been exhausted by the subject. Given two subjects with the same number of questions required to solution, this index is a measure of the stereotypy of response. If one analyzes the breakdown of total questions into separate items, one finds a marked inhomogeneity of distribution. Thus, of 174 questions asked by one subject in a performance on the first problem, 85 consisted of or included the same item. It might be of interest to plot the rank order of the frequency with which a given item appears versus its actual frequency in order to see whether the logarithmic relation described by Zipf for word frequency distributions can be demonstrated with a symbolic language of this abstract type. One might also tabulate response frequencies with the end of categorizing "popular" and "original" responses. These analyses have not been performed.

A frequency analysis of the sort described above, were the logarithmic prediction upheld, would give us insight into the mode of organization of the process in a descriptive sense. As this variable is used at present, it is more a product than a process index. Examination of the correlation between this measure and Actual Redundancy (see following), which is a process index, shows that the two are almost completely independent. This is an unexpected finding. One would normally expect to find that as stereotypy of response increases, redundancy increases. This apparent contradiction is probably due to the existence of two sources of low values for this measure: first, the incisive performance requires only a few items to be elicited; second, the disorganized performance may be highly stereotyped, or may be highly exploratory but undirected. Given two persons who required the same number of questions for solution, one might expect the one with a lower Exhaustiveness measure to demonstrate less flexibility conceptually, provided that both performances were relatively long. The data suggest that educational level and specialized training affect this variable very slightly. It is of interest that members of all the groups consistently select two items in each problem which can be inferred by inspection a priori.

2. ACTUAL REDUNDANCY—the total number of *inferrable* questions, as defined in the preceding discussion, divided by the total number of questions asked. (Data for this variable are presented in Table 10.)

Correlations. Actual Redundancy correlates with Questions (problem 1) .72, (problem 2) .77, with Rate .57, with Effort .72, with Exhaustiveness of Inquiry .10, with Percentage of Nonlinearity .24, with Mixture of Modes (see following) −.24, with Predominant Mode (see following) .08, with Frequency of Change of Approach (see following) .49, with Relative A-S Shift (see following) −.39, with Absolute Inferential Lag

(see following) .59, with Absolute Synthetic Lag (see following) Explicit −.41 and Implicit .47, with ACE "Q" scale −.24, ACE "L" scale −.14, and Total Analysis −.46.

Discussion: Actual Redundancy is a process variable, in the sense that it describes the logical efficiency of the problem-solving performance as a whole by examining it step by step. As pointed out in the theoretical discussion in which this measure was developed (see Section F, paragraph 5), one expects that ordering as well as the actual questions asked will determine the rate of acquisition of information. In support of this, note that the correlation between Questions and this variable is .72 and .77 for the two problems used. This indicates that about 48% of the variance in Actual Redundancy on the easier problem, and 41% on the more difficult problem, must be attributed to the *ordering* of the questions asked. This gives some idea of the importance of sequential analysis of the sort here used in proper evaluation of the quality of a process.

Further inspection of the correlations presented above suggests that fast workers tend to be more redundant. Redundancy of this sort is slightly related to a tendency to carry out analysis separate from synthesis, as we see from the correlation with Mixture of Modes. It is relatively independent of whether the performance is predominantly analytic or synthetic, as we see from the correlation of .08 with Predominant Mode. The correlation with Frequency of Change of Approach shows that Redundancy is related to a large amount of shift from synthesis back to analysis, and vice versa. As Redundancy goes up, the change from the analytic phase to the synthetic phase comes relatively earlier in the performance; this change occurs more prematurely in the less redundant workers, with respect to the achievement of the information needed for solution, as we see from the correlations with Relative A-S Shift and Absolute Synthetic Lag. Finally, as Redundancy increases, the lag in inference increases, as seen from the correlation with Absolute Inferential Lag.

The correlation of this measure from problem 1 to problem 2 is .60.

Examination of the data above shows that this index is somewhat dependent on educational level and specialized training or interests of a technical sort. The first effect may be seen by comparing the more educated groups (1, 3, and 4) with the other groups. Note, however, that group 5, which is a college group, is as effective as group 4, which is a Ph.D. group, with respect to this index. Note the relatively uniform increment in Redundancy which all groups display as the difficulty of the problem increases.

Note also that the correlation of this index with the ACE scales is not high, while there is a correlation significant at the .01 confidence level between this index and the derived score for Total Analysis on all achievement tests for which data are available.

3. NASI Point (Implicit)—the point in the performance where the subject obtains the information which is necessary and sufficient to enable solution of the problem, provided that maximum inferential use is validly made of the data on hand. This point is stated in two ways: (*a*) Absolute—the number of questions asked up to and including the point; and (*b*) Relative—as a percentage of the total number of questions. (Data for these variables are presented in Table 11.)

Correlations. None computed.

Discussion. This measure serves as an index of the end of the first phase of inquiry in the problem-solving process as observed using the PSI. Once the NASI (Implicit) has been achieved by the subject, he is theoretically able to solve the problem. That is, he has had the opportunity to learn the required relationships on which a solution must be based. Note that for all groups this point is achieved at about the same number of questions in both problems, with the excep-

TABLE 10

Actual Redundancy

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	.79	.70	.83	.61	.79	.79	.89
Problem 2	.86	.78	.92	.69	.86	.83	.94
Difference	−.07	−.08	−.09	−.08	−.07	−.04	−.05

TABLE 11

NASI POINT (IMPLICIT)

Problem	Absolute						
	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	14.6	14.2	17.9	12.2	16.1	12.2	12.8
Problem 2	15.4	12.7	18.3	10.6	14.6	11.4	20.3
Difference	−.8	1.5	−.4	1.6	1.5	.8	−7.5
	Relative						
Problem 1	.40	.47	.41	.56	.39	.41	.24
Problem 2	.27	.34	.26	.40	.27	.26	.18
Difference	.13	.13	.15	.16	.12	.15	.06

tion of group 6. The slight increase in difficulty involved when the second problem is presented has little consequence on the location of the NASI (Implicit), Absolute. There is evidence that educational level or special training are effective in slightly earlier achievement of the NASI (Implicit).

The NASI (Implicit) Relative tells us the percentage of the total process required for the achievement of the NASI (Implicit). Note that as educational level or special technical training increases, the relative location of the NASI (Implicit) is later in the performance. Since we know from the absolute location that there is little difference in the numerator of the fraction which determines the relative value, the difference which we observe here is mostly due to the denominator, i.e., total questions in the various performances. Note that the effect of the increase in difficulty is to shift the location of the NASI (Implicit) point about 13% earlier in the performance, except for group 6.

4. NASI POINT (EXPLICIT)—the point in the performance where the subject obtains the information which is necessary and sufficient to enable solution of the problem without the necessity of inferring any relationships involved. (That is, the relevant items have been overtly elicited from the PSI.) This point is stated in two ways: (*a*) Absolute—the number of questions asked up to and including the point; and (*b*) Relative—as a percentage of the total number of questions. (Data for these variables are presented in Table 12.)

TABLE 12

NASI POINT (EXPLICIT)

Problem	Absolute						
	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	46	31	53	21	40	53	59
Problem 2	66	38	71	24	50	68	110
Difference	−20	−7	−18	−3	−10	−15	−51
	Relative						
Problem 1	.87	.92	.84	.94	.90	.88	.80
Problem 2	.74	.81	.68	.77	.86	.73	.72
Difference	.13	.11	.16	.17	.04	.15	.08

Correlations: None computed.

Discussion. This measure serves as an index of the end of the second phase of inquiry in the problem-solving process as observed using the PSI. We conceive of the first phase, ending with NASI (Implicit), as an interval in which basic structuring of the problem occurs. At the end of the first phase, the individual has elicited the information which is required minimally for solution to be achieved. Almost invariably, solution based on such inferences is not forthcoming. Instead, there intervenes a period of further inquiry which culminates in the NASI (Explicit), in which the material which might have been inferred at the end of the first phase is overtly elicited. This second phase we tend to conceptualize as a reassurance period in which verification takes place of the inferences which can be drawn after the first phase. Of course, the fact that inferences *can* be drawn does not mean that the subject actually *does* draw them in the second phase. It might well be that the activity in the second phase is frequently required for perception of relationships to become accurate. Yet it is of interest that one frequently observes subjects who verbalize as they work in the second phase, and it becomes apparent that inferences are being tested. Further, very seldom does one observe a performance in which the subject does *not* go ahead after achieving the NASI (Implicit) to achieve it explicitly. Subjects who change to synthetic behavior, as defined later, before the NASI (Explicit) is achieved, almost invariably revert to analysis subsequently until the NASI is achieved explicitly.

Examination of the data shows that a marked differential exists on the absolute index between our various groups. This differential is such as to support the interpretation that both educational level and specialized technical training or interests contribute to the ease with which the subject achieves the NASI explicitly. Since the various groups end the first phase with the implicit achievement of the NASI almost identically, this difference in the location of the NASI (Explicit) point must be due to the variable length of the interval *between* the two points. This will be discussed in the next section.

Note that groups with higher educational levels and greater interest or training in specialized technical areas tend to achieve the NASI explicitly relatively later in the over-all performance. This implies that these groups have less difficulty in combining and applying the results of analysis than groups with less education or technical sophistication.

Finally, note that the decrement in the Relative NASI (Explicit) as the difficulty of the problem is increased is almost identical with that for NASI (Implicit), with the exception of group 4. This means that as the difficulty of the problem is increased there is a decrease in the portion of the problem which is occupied by the first and second phases of analysis. Therefore it follows that there must be a compensatory increase in the relative portion of the process occupied by activities other than analysis of this sort, as difficulty increases. In other words, the effect of increasing the difficulty of the problem is not to cause relatively more analytic trouble, but rather to cause relatively more synthetic difficulty.

5. Inferential Lag—the number of questions which intervene between the achievement of the NASI implicitly and explicitly. The Inferential Lag can be stated in two ways: (*a*) Absolute—the actual number of questions in the interval between the two NASI points; and (*b*) Relative—the number of questions in the interval divided by the total number of questions required for solution. (Data for these variables are presented in Table 13.)

Correlations: Absolute Inferential Lag correlates with Effort .77 and with Actual Redundancy .59. Relative Inferential Lag correlates with Effort .30.

Discussion. The two Inferential Lag indices are measures of the quality of the inferences drawn by the subject in the course of the process. They are indicative of the failure of the subject to infer properly from the data he has gathered or of his need for reassurance about inferences which he has made, which he tests by making the conclusion overt rather than by attempting to utilize the conclusions in attempted solutions. The fact that the NASI (Explicit) occurs in a performance once the NASI have been achieved implicitly is evidence that conclusions which could have been drawn either have not been drawn or have not been accepted. In the interval, the subject accumulates direct evidence which provides the NASI without inference. Since this material is redundant, it is obvious that this index should correlate well with our redundancy index for the whole performance, as it does. It is also obvious that difficulty in the second phase of the process, as measured by the lag, will correlate significantly with our index of over-all difficulty—Effort.

Examination of the data for the absolute index suggests strongly that both educational level and the amount of training or interest in technical fields are related to the size of the inferential

TABLE 13

INFERENTIAL LAG

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
				Absolute			
Problem 1	40.0	17.2	34.7	9.8	24.0	40.7	46.3
Problem 2	62.7	25.5	34.0	13.7	36.5	55.5	89.0
Difference	−22.7	−8.3	0.7	−3.9	−12.5	−14.8	−43.7
				Relative			
Problem 1	.52	.45	.44	.38	.50	.46	.56
Problem 2	.50	.48	.40	.37	.59	.48	.54
Difference	.02	−.03	.04	.01	−.09	−.02	.02

lag. Note, however, that the consequence of increasing difficulty of the problem is to lengthen this interval for all groups except the first-year college group, group 2. The difference between groups 1, 3, and 4—the more educated groups—and the less educated groups is more perceptible on the easier problem. The difference between groups 3 and 5 and their less technically sophisticated counterparts at equivalent educational levels tends to be more marked on the more difficult problem. This suggests that as the complexity of the task increases, education as such is less advantageous than technical sophistication.

Note that the more technically sophisticated groups devote relatively less of the over-all PSI performance to the inferential lag, while this distinction does not appear in favor of the more educated group compared with the first-year group (1 vs. 3). Finally, let us point out that the effect of increasing the difficulty of the task on the proportion of the over-all performance spent in the inferential lag is remarkably little. This suggests that increase in difficulty is not followed by an increase in the relative difficulty of inference. However, the correlations between performance on the two problems are .30 for the absolute index and —.12 for the relative index. The first correlation is significant at the .05 level. This indicates that the effect of increasing the difficulty of the problem is not the same on all individuals. Again, it should be pointed out that the correlations just given do not indicate reliability, since this is a measure of organization, and one has no reason to expect similar organization of process for two tasks which differ in the severity of the demands made on the individual. Tasks of equated difficulty must be used to obtain such a measure of reliability and this has not been done.

G. *Approach Variables*

A graph is drawn of the duration of the average question in each minute plotted versus time, to which we will subsequently refer as the Analysis-Synthesis (A-S) Graph (see Section I). To draw this graph, we sum the series of manipulations of the apparatus which occur in each minute, divide by the number of questions asked in that minute, and plot the result versus that minute.

We assume that the following distinction can be made between two modes of approach to the problem. Direct questions about particular propositions or relationships can be asked with very short series of manipulations. Such short manipulative combinations we term as of the *analytic* mode. When the subject attempts to utilize relationships which he has analyzed to achieve solution of the problem, combining them and causing interactions to occur between them, the series of manipulations composing a question becomes longer. Such long manipulative combinations we term the *synthetic* mode.

Rigorously, the A-S graph is analyzed as follows: A time interval is classed as *analytic* if there are less than two ordered

related manipulations (steps) per average question in that interval, *or* if there are exactly two steps per question in that interval and the interval is bracketed on either side by intervals with less than two steps per question, *or* if there are exactly two steps për question in that interval and only one adjacent interval has less than two steps per question and only one question was asked in the relevant interval. A time interval is classed as *synthetic* if there are more than two steps per question in that interval, *or* if there are exactly two steps per question in the interval and it is bracketed by two intervals each of which contains more than two steps per question, *or* if there are exactly two steps per question and only one adjacent interval has more than two steps per question and more than one question was asked in the relevant interval. Note that this analysis could be carried out question-by-question instead of minute-by-minute. Such an analysis would be more precise but much more laborious and has not been extensively pursued.

The performance can in this fashion be characterized as a period of so many minutes of analysis (A) followed by a period of so many minutes of synthesis (S) followed by . . . etc. Consider a performance in which five minutes of A is followed by three minutes of S is followed by five minutes of A is followed by seven minutes of S. We write this as 5A3S5A7S. The total time of this performance is 5 plus 3 plus 5 plus 7, or 20 minutes. There was a total of 10 minutes spent in the analytic mode, or 50%. Ten minutes were spent in the synthetic mode, or 50%. (Note that if this analysis were done on the basis of a question-by-question analysis, the straight line phenomenon described in the section on Output Graphs would result in approximately the same distribution of analysis and synthesis.)

It can be shown that, at any point in a performance, the following is true: the percentage of analytic before that point *minus* the percentage of synthetic before that point *is equal* to the percentage of synthetic after that point *minus* the percentage of analytic after that point. We are now in a position to define a number of variables.

1. Analytic-Synthetic Shift Point (A-S Shift)—that point in the performance where the percentage of Analytic before minus the percentage of Synthetic before is a maximum for the whole performance. This defines a unique point which separates the PSI performance into a predominantly analytic phase before the point and a predominantly synthetic phase after the point. This index can be stated as (*a*) Absolute, and (*b*) Relative. (Data for these variables are presented in Table 14.)

Correlations. The Relative A-S Shift correlates with Effort −.59, and with Actual Redundancy −.39.

Discussion. The Analytic-Synthetic Shift Point, together with the indices to be presented in the remainder of the section on approach variables, permits us to measure the extent to which the PSI performance does consist of separate and separable phases of information gathering and information application. It will be seen in what follows that such a separation is possible, and that examination of the relationship between the state of information possessed by the subject and the transition from inquiry to application is quite informative. Note that as the shift comes *relatively* earlier in the performance, both Effort and Redundancy tend to increase.

Examination of the data above suggests that more educated groups tend to shift earlier than less educated groups, and that more technically sophisticated groups tend to shift earlier than less sophisticated groups. Note that these differences do not hold for the relative location of the shift point. It is exceedingly interesting that the change in the absolute shift point is so little when the difficulty of the problem is increased.

TABLE 14

ANALYTIC-SYNTHETIC SHIFT POINT (A-S SHIFT)

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
	Absolute						
Problem 1	46.3	26.3	49.0	18.4	33.7	28.0	62.5
Problem 2	47.9	28.4	48.5	24.2	32.3	35.0	66.5
Difference	−1.6	−2.1	.5	−5.8	1.4	−7.0	−4.0
	Relative						
Problem 1	.70	.75	.77	.79	.71	.65	.78
Problem 2	.55	.59	.73	.64	.54	.58	.50
Difference	.15	.16	.04	.15	.17	.07	.28

It suggests that after a certain amount of inquiry there is a tendency for the subject to shift into synthesis, "ready or not." However, we note from inspection of the data for the relative index that the shift occurs appreciably earlier for all groups on the second problem. From this we conclude that the consequence of increasing the difficulty of the task is to increase the proportion of the process which is in the synthetic phase, rather than to increase analytic aspects of the performance. That is, as the difficulty of the problem goes up, about the same effort appears to be required to acquire the necessary information but it becomes more difficult to reconcile the constraints which this information imposes.

The correlation between the two performances for the absolute index is .13, and for the relative index is .33. This suggests that there is a reorganization of the internal coherence of the PSI performance as difficulty increases.

2. PREDOMINANT MODE—total number of Analytic intervals divided by the total number of Synthetic intervals. (Data for this variable are presented in Table 15.)

Correlations. Predominant Mode correlates .08 with Actual Redundancy and −.18 with Rate.

Discussion. This index is a measure of the relative amounts of analysis and synthesis of which a PSI performance is composed. We see from the correlation of .08 with Actual Redundancy that this index is relatively independent of the efficiency of the performance, in terms of economy of inquiry; the same holds true of Rate. The data show that on the easier problem there is little relation between educational level and the A-S ratio. The technically sophisticated groups tend to engage in somewhat less analysis relative to synthesis than their counterparts at the same educational levels. On the more difficult problem neither educational level nor special technical skill or interest appears to be related to the A-S ratio.

Notice that all groups except group 5 show a marked decrease in analysis relative to synthesis as the difficulty of the problem increases. This supports the inference which was earlier drawn that increased difficulty has the effect of causing greater synthetic effort rather than analytic effort. That is, the manifestation of increased difficulty in the PSI problems here used is to bring about a shift toward more synthetic en-

TABLE 15

PREDOMINANT MODE

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	6.3	5.6	5.0	5.0	6.2	4.1	9.7
Problem 2	3.3	1.7	1.8	1.7	1.7	3.9	5.3
Difference	3.0	3.9	3.2	3.3	4.5	.2	4.4

TABLE 16

MIXTURE OF MODES

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	.79	.85	.76	.91	.80	.75	.89
Problem 2	.66	.64	.81	.63	.65	.76	.53
Difference	.13	.21	−.05	.28	.15	−.01	.36

deavor in the process, because of the complexity of the constraints which must be reconciled. This should be considered in view of the fact that the two problems, as described earlier, do not differ in the total size of the pool of information items which stipulates the problems, but differ in the number of coincidences which must be achieved using these relationships. The reader may recall that the more difficult problem requires the subject to achieve four coincidences, the easier only three. Presumably, it is to this difference that we can attribute the above shift toward synthesis.

The correlation between the two performances for this index is .70. Thus, even though there is a shift in the internal organization of the process as the difficulty is increased, the way in which an individual distributes his efforts between the acquisition of information and the utilization of that information remains fairly characteristic, in the sense that one can relate the new organization to the old.

3. MIXTURE OF MODES—the percentage of the total number of analytic intervals in the performance which is located before the shift point minus the percentage of the total synthetic intervals in the performance located before the shift point. (Data for this variable are presented in Table 16.)

Correlations: Mixture of Modes correlates with Actual Redundancy −.24 and with Rate .08.

Discussion. This index measures the extent to which PSI performance can be separated into distinct analytic and synthetic phases. It is apparent from an inspection of the data that these two phases exist relatively separate from one another in both problems, although the definiteness of the separateness is greater on the more simple problem. Factors such as education or specialized technical training or interest do not appear to relate to the extent of this separation in a clear fashion. Note, however, that the definite separation decreases more over the more educated groups than over their counterparts as the difficulty of the problem increases, with the exception of group 6. The meaning of this differential is not clear.

The low correlation of this index with Actual Redundancy and with Rate suggests that we have in the Approach Variables access to a set of dimensions relatively orthogonal to those which span the space of Work Variables and Information Variables.

In summary, then, this variable enables us to distinguish the intermingling of two subsidiary processes which are combined in PSI performance. The high values which are obtained, showing relatively clean separation of the process into two phases, indicates that the Analytic-Synthetic Shift Point, rather than being an artificial conceptual device, is a concept which has real functional utility, enabling the separation of two aspects of process which are not homogeneously distributed in the performance.

The correlation for this variable between the two performances is −.39. We note from the data above that the greater the separation between the two phases on the easier problem, the greater will be the changes as the difficulty of the problem is increased. Perhaps individuals whose performance on the easy problem is composed of exceedingly distinct phases are more venturesome on the second problem. Although the reason for this peculiar relationship is far from obvious, the correlation is significant at the .01 level (negative sign and all).

4. FREQUENCY OF CHANGE OF APPROACH—the total number of inversions from synthesis back to analysis which occur in the performance. (Data for this variable are presented in Table 17.)

Correlations. This index correlates with Effort .59, and with Actual Redundancy .49.

Discussion. The previous index, Mixture of Modes, gave us a measure of the separation of the two phases, but did not enable us to dis-

TABLE 17
FREQUENCY OF CHANGE OF APPROACH

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	3.6	2.7	3.3	2.0	3.3	3.4	2.6
Problem 2	7.8	7.0	5.7	6.5	7.4	4.8	10.8
Difference	−4.2	−5.3	−2.4	−4.5	−4.1	−1.4	−8.2

criminate between the performance which interrupted an analytic phase with a prolonged synthetic interjection and one which shifted back and forth repeatedly. The present index tells us which case we are dealing with. It measures the number of times that a subject attempts to synthesize and feels forced to revert to analysis for more information. We can see by comparing the correlations with Redundancy of this and the previous index that mixture of modes per se does not relate to a higher redundancy, but that frequent and premature attempts at synthesis as measured by frequency of change of approach do so relate.

Neither the frequency itself nor the change in frequency with greater difficulty appears to be highly related to either educational level or specialized technical training or interest. We note that all groups change more frequently on the more difficult problem.

The correlation between the two performances for this index is .04. This may be a consequence of the much greater spread for this variable on the more difficult problem.

5. Synthetic Lag (Implicit)—the Absolute A-S Shift Point minus the Absolute NASI Point (Implicit). This index is stated as (*a*) Absolute and (*b*) Relative. (Data for these variables are presented in Table 18.)

Correlations. Absolute Synthetic Lag (Implicit) correlates with Actual Redundancy .47.

Discussion. This is a measure of the failure of the subject to shift from analysis to synthesis as soon as the sum of inferences which can be validly drawn from the information he has gathered is sufficient for solution to occur. Negative implicit synthetic lag occurs very rarely.

The most impressive fact about this index is its remarkably small change when the difficulty of the problem is increased. Neither the absolute nor the relative versions of this index show large changes with difficulty, except for group 6 in the relative category. These data show that shift tends to occur after a characteristic interval, independent of the objective difficulty presented by the problem, if one considers group averages. This interval is shorter for the educated groups than for the less edu-

TABLE 18
SYNTHETIC LAG (IMPLICIT)

Problem	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
	Absolute						
Problem 1	29.0	12.6	31.4	7.2	17.7	15.8	49.5
Problem 2	32.4	16.2	29.7	13.6	18.6	22.7	47.2
Difference	−3.4	−3.6	1.7	−6.4	−.9	−6.9	2.3
	Relative						
Problem 1	.38	.27	.35	.23	.31	.24	.54
Problem 2	.33	.25	.42	.23	.27	.32	.33
Difference	.05	.02	−.07	.00	.04	−.08	.21

TABLE 19

SYNTHETIC LAG (EXPLICIT)

Problem	Absolute						
	All Groups	Group 1 (Ph.D.)	Group 2 (1st yr. College)	Group 3 (Nat. Sci. Ph.D.)	Group 4 (Other Ph.D.)	Group 5 (Nat. Sci. College)	Group 6 (Other College)
Problem 1	−11.0	−4.6	−3.4	−2.6	−6.4	−24.9	−3.2
Problem 2	−30.0	−9.4	−4.2	−.1	−17.9	−32.7	−42.7
Difference	19.0	4.8	.8	−2.5	11.5	7.8	39.5
	Relative						
Problem 1	−.14	−.17	−.08	−.15	−.19	−.22	−.04
Problem 2	−.17	−.23	−.03	−.00	−.30	−.16	−.22
Difference	+.03	+.06	−.05	−.15	+.11	−.06	+.18

cated groups, and shorter for the technically more sophisticated groups than for the others. This appears to hold true for both versions of the index, but is less clear for the relative measure.

The correlation between this variable and Actual Redundancy, significant at the .01 level, is to be expected. Most of what occurs in this interval acts to increase redundancy.

The correlation between the two performances for this variable is .01 for the Absolute version and .09 for the Relative. The author is at a loss as to the explanation for this low correlation in view of the small group changes.

6. SYNTHETIC LAG (EXPLICIT)—the Absolute A-S Shift Point minus the Absolute NASI Point (Explicit). This index is stated both as (*a*) Absolute and (*b*) Relative. (Data for these variables are presented in Table 19.)

Correlations. Absolute Synthetic Lag (Explicit) correlates with Actual Redundancy −.41.

Discussion. This index is a measure of the failure of the subject to shift to synthesis as soon as the explicit yield from analysis is sufficient for solution of the problem to occur. The following considerations are relevant to an evaluation of this index:

1. If the A-S shift occurs before the NASI Point (Explicit) is achieved, the synthetic lag is negative. Such a premature shift may be adaptive if the subject does not find it subsequently necessary to return to analysis for further items of information or for the NASI (Explicit) before achieving solution of the problem, since one then concludes that the NASI (Implicit) was the basis of operation. If the subject finds it necessary to return to analysis (inversion) after the shift occurs, the premature shift appears to be inappropriate, indicating overly flexible performance.

2. If the A-S Shift occurs after the NASI is explicitly achieved, the Synthetic Lag is positive, and its magnitude is an indication of the extent to which further inquiry is required by the subject before realization that the information already acquired constitutes an adequate basis for solution of the problem.

3. Whether the Synthetic Lag is positive or negative, its absolute magnitude is a measure of the extent to which shift from analysis to synthesis is inappropriate, provided that the NASI (Explicit) appears in the performance.

For all groups, the synthetic lag is negative on both problems. Very seldom does one observe a performance in which it is positive. The extent to which the shift is premature does not appear to depend on educational level or specialized technical training or interests. The change in the magnitude of the lag which ensues as the level of difficulty of the problem is increased also appears to be independent of these factors. In general, the size of the lag increases as difficulty increases; that is, shift tends to occur more prematurely, except for group 3.

Most of the above discussion also applies to the relative version of the index. Note, however, that as difficulty increases the lag for group 3 becomes relatively smaller, while that for group 4 increases. Similarly, the lag for group 5 decreases while that for group 6 increases. Since educational level is relatively constant in each of these two comparisons, perhaps the greater technical sophistication of group 3 and 5 explains the fact that groups 4 and 6, in contrast to 3

and 5, shift relatively even more prematurely on the difficult problem. The correlation of —.41 between the Absolute Synthetic Lag and Actual Redundancy, which is significant at the .01 level, means that as the lag becomes more *negative*, redundancy becomes larger.

Probably a small negative lag would be associated with a less redundant performance than a small positive lag, since a positive lag precludes good guessing. One would also expect large lags, whether negative or positive, to be indicative of higher redundancy. These statements, of course, are speculative, since we have so little data on positive lag. The correlation between the two performances is .49 for the Absolute index and .40 for the Relative index. The tendency to shift to synthesis before analysis is completed appears to be fairly consistent for a given individual. Both correlations are significant at the .01 level.

H. Mean Values and Standard Deviations for All Variables

The mean values and standard deviations obtained on the two problems by our full population are summarized in Table 20. It will be noted that not all the variables just defined are presented. Some of these variables are used in the calculation of others and for no other purpose. The correlation between performance on the two problems is also presented for each variable. Table 20 is then a summary of the data which has been presented under the heading "All Groups."

Once more let us emphasize that this set of data should not be regarded as an accurate indication of the consistency of PSI performance for the individual. Some aspects of PSI behavior are markedly changed when problem difficulty increases, perhaps due to the necessity for reorganization of process when the severity of the demands imposed by the

TABLE 20

MEANS, STANDARD DEVIATIONS, AND "RELIABILITY" OF VARIABLES

(N=45, r=.297 for p=.05, r=.380 for p=.01)

Variable	Problem 1			Problem 2	
	Mean	*SD*	r_{12}	Mean	*SD*
Work Variables					
Time	23.8	16.8	.462	44.5	21.9
Questions	69.1	53.4	.462	110.0	86.8
Complexity	1.71	.23	.112	2.30	.69
Rate	3.01	1.25	.844	2.27	1.15
Pauses	.09	.16	.758	.18	.20
Changes of Set	1.5	1.2	.506	2.5	1.5
Percentage of Nonlinearity	12.6	11.8	.022	22.5	18.2
Effort	379.2	521.5	.424	814.5	733.5
Information Variables					
Exhaustiveness of Inquiry	13.6	2.5	.184	13.0	2.5
Actual Redundancy	.79	.10	.597	.86	.09
Inferential Lag (Absolute)	40.0	45.7	.297	62.7	74.9
Inferential Lag (Relative)	.52	.26	−.121	.50	.32
Approach Variables					
Analytic-Synthetic Shift (Abs.)	46.3	39.9	.126	47.9	35.0
Analytic-Synthetic Shift (Rel.)	.70	.21	.334	.55	.25
Synthetic Lag, Explicit (Abs.)	−11.0	35.8	.487	−30.0	68.1
Synthetic Lag, Explicit (Rel.)	−.14	.26	.398	−.17	.38
Synthetic Lag, Implicit (Abs.)	29.0	32.8	.003	32.4	34.1
Synthetic Lag, Implicit (Rel.)	.38	.27	.086	.33	.25
Predominant Mode (A/S)	6.32	7.58	.696	3.29	7.77
Frequency of Change of Approach	3.6	2.9	.042	7.8	5.8
Mixture of Modes (A minus S before shift)	.79	.20	−.387	.66	.24

task on the subject increases. To determine reliability, it will be necessary to construct two problems of more comparable structure than the two here used, and then to analyze the results obtained from the administration of these. While it is important to ascertain such data, this was not possible in this exploratory work. This shortcoming is to be regretted. We gain here, in addition to an insight into the effects of increasing difficulty, only an approximate idea of the consistency of an individual's behavior on repeated exposure to the PSI. In view of the known differences between the two problems, we might assume that reliability figures properly obtained should be at least as high as the correlations so far presented.

I. Summary of Discussion of Variables

Now that we have presented the full set of variables so far developed for the PSI, we may gain an overview of the analysis which they enable. In Fig. 5, we present a combined Output and Analysis-Synthesis Graph, on which a number of the variables defined and discussed in the previous pages are illustrated. Variables from all three areas of PSI performance are represented in this illustration.

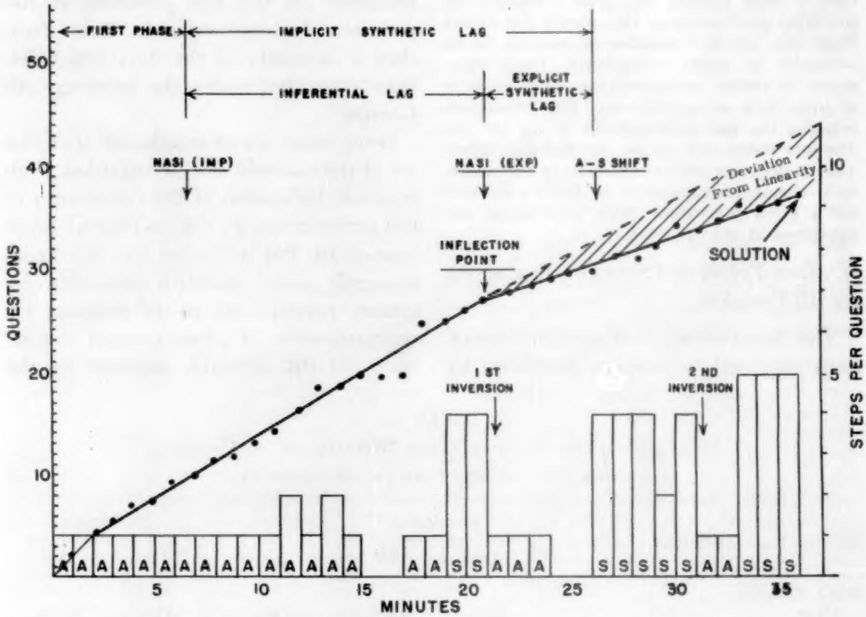


Fig. 5: Output and Analysis-Synthesis Graph. This illustrates the salient features on which PSI scoring is based.

The set of variables which we have defined gives information about a number of different aspects of the problem-solving process. Some of these variables appear to be sensitive to personality as well as to "cognitive" factors manifested during the performance. Such variables as inferential lag, synthetic lag, frequency of inversion, and even actual redundancy, to name a few, one would expect to be particularly susceptible to such factors as the effects of test

anxiety. A thorough analysis of the extent to which the variables listed above, which demonstrably describe characteristic and important parts of a person's problem-solving behavior, at least on the PSI and perhaps in other tasks, can serve as indices of personality organization is beyond the scope of this paper. Some definitely preliminary results relevant to these considerations will be mentioned in Section IV following, together with a brief and also preliminary survey of the relationship between some aspects of PSI performance and other less abstract cognitive tasks. The entire population of this study has been recalled for intensive psychological evaluation by Mr. Sidney Blatt, of the Department of Psychology of the University of Chicago, working in conjunction with Dr. Morris Stein of the same department. These same workers are also currently engaged in research to investigate the relationship between creativity in research chemists and PSI performance. Results from these researches will eventually be forthcoming from their laboratory.

On the basis of the work already carried out, it appears clear that much more is involved in PSI performance than cognitive factors alone. Examination of the extent to which stress will change characteristic performance on some of these variables, the extent to which others are relatively invariant across many conditions, correlative studies between this and other behavioral measures—all must be used before one can specify the dependence of PSI behavior on perceptual, personality, and cognitive elements.[8]

There is an obvious inadequacy to the variables so far developed, if our goal is to understand the manner in which the problem-solving process is generated by the interaction of an individual with a problem. Initially, the subject has a choice between a large number of possible moves, with relatively little basis for discriminating between these as to utility. After the first questions asked of the PSI, the information which is forthcoming restructures the probabilities of various questions being asked next, so that instead of approximately equal probabilities for all questions, we now have a more probable and a less probable subset. The next choice is made, in a logical performance, from the more probable subset, with a consequent alteration of probabilities for all moves in accordance with the information derived from the two moves so far completed.

Ideally, we would like to be able to state why the *n*th step follows the *n—first* step, and in what way it derives from and relates to the previous elements in the process. This requires that we attempt to relate the state of the information possessed by the subject to the next step performed by the subject. We would hope ultimately to be able to include in the description of state those parameters of personality and perception and cognition which will affect choices between steps of approximately equal probabilities based on information considerations alone. If this is our goal ultimately —to understand the dynamics of the process through which each step is generated—we must acquire both an adequate description of state and a knowledge of the function by which the transition probability to the next step is determined. We have here stated a problem which is recognized today as that of analyzing a stochastic process with nonindependent transition probabilities. Our goal is to learn the function which stipulates the manner in which transitions from element to element of the process take place. It is hoped that the PSI will lend itself to this sort of analysis. The present author feels, however, that some of the variables so far developed represent largely static abstractions rather than dynamic functions, and will require much improvement before we can approach the above-mentioned goal.

III. PSI Behavior

A. General Description

A general description of problem-solving behavior as demonstrated on the Problem-Solving and Information Apparatus (PSI) can now be presented. This qualitative sketch represents a "typical" performance, to which numerous exceptions can be found in the data.

The starting point chosen by the subject for his initial attempts to structure the PSI

[8] To save printing costs, a discussion of the effects on PSI performance of education per se, education in a particular discipline, and interest in a particular discipline before specialized training, as well as development of a performance profile, an evaluation of our set of variables as comparative indicators, definition and discussion of a utility index, examination of the consequences of increased difficulty on our variables, and a comparison of PSI performance of a small group of college students with achievement test scores, college records, and personality indices, have been deposited with the American Documentation Institute. Order Document No. 5359 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., remitting in advance $2.25 for 35 mm. microfilm or $5.00 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

situation is usually at the "output" end of the problem, and consists of a rapid and generally exhaustive investigation of the immediate antecedents of the output. Very seldom, one encounters an individual whose initial procedure is to permute the "input" elements in a systematic fashion. The problems are designed to make this approach difficult and tedious, and it usually gives way to the more systematic output analysis described above.

As the antecedent conditions for the output are determined, the rate of asking questions diminishes sharply and a slower and steady rate is established and maintained. At this point, many subjects commence verbalization which continues until solution is achieved, and start to take notes about the events which they observe. Verbalization, when it occurs, is of great utility in aiding understanding of the basis for certain aspects of the performance. Pauses occur during which notes and the apparatus are scrutinized and "self-discussed" aloud. These pauses are frequently followed by a transient acceleration in rate which persists for a time proportional to the length of the pause. The rate of reaching decisions appears to be very stable and characteristic of the individual, and is manifested by a steady linear increase in the cumulative question curve as drawn on the Output Graph. Some system of notation is adopted early in the performance, which serves as a shorthand rather than as a conceptual aid.

One might represent the progress of an individual through the PSI problems as a series of reformulations of the problem, each time focusing on different aspects of the many relationships involved. After the immediate antecedents of the output are discovered, the problem is restructured, and becomes in effect the problem of achieving these antecedents. Since there are several antecedents, they constitute what are in a sense subproblems. These are solved separately at first, and the result of the first restructuring tends to be concentration on the task of learning how to achieve these antecedents by using the permitted input elements. However, at this phase, the various antecedents are considered separately, and their achievement tends to be viewed as separate problems. The PSI problems used in the experiments reported here are so built that simple combination of the conditions which will achieve the antecedents separately results in an interaction between these conditions which precludes the solution from being achieved.

Recognition that the interaction between the simple solutions is such as to introduce a new factor into the constraints previously recognized as relevant in the problem is the prelude to a revision and reformulation of the problem as perceived by the individual. It is at this point that most subjects encounter their first substantial difficulty. First, one must analyze the interaction in order to determine the manner by which solution is prevented. This phase of the problem is often accompanied by long and amusingly emotional verbalized arguments which tend to be syllogistic. Ultimately, the subject discovers the source of the blockage, and this discovery, even in staid adults, is greeted with great pleasure and satisfaction.

The discovery of the aspect of the interaction between the separate solutions to the subproblems initiates the final phase of the process, in which the subject devises a way to block the interaction. This consists of analyzing relationships in order to learn how to bring about a result which will prevent another result from occurring. This is the most serious source of difficulty in completion of analysis.

When information is acquired by the subject which might be expected to change set—in that it affords the opportunity to achieve closure of one of the aspects of the problem—an inflection point is often observed in the Output Graph; that is, the rate at which questions are asked changes to a new and also steady rate. This phenomenon is particularly of interest when the achievement of the necessary and sufficient information (NASI) for solution of the problem occurs implicitly; that is, when the information from which solution can be inferred is in the possession of the subject. One frequently observes a change in rate at this point, even though the subject continues analysis until he achieves these items of information (the NASI) overtly. This indicates that the significance of the information which was on hand at the point where the inflection occurred was not effectively incorporated by the subject even though it did cause a change of set. Surprisingly, the overwhelming majority of our sample went ahead to obtain the NASI explicitly before they would attempt to produce a solution to the problem by a sustained combinatory approach. One tends to conclude that inference once drawn is confirmed by direct inquiry instead of by an attempt to apply the results of inference, thus testing validity implicitly. The inferential lag may be an index of the amount of reassurance needed before an inferred conclusion is accepted and translated into action.

Although the initial investigation of the antecedent conditions of the output is usually systematic, there is generally not a sustained methodical tracing-back of these antecedents through the network to the input end. A particular relationship is focused upon and clarified, and then attention is frequently directed to another aspect of the problem which is unconnected with the relationship previously investi-

gated. There is a tendency for the completion of what might be termed "blocks" of analysis, which relate to the subproblems referred to earlier. These blocks involve groups of related or contingent relationships. However, such block analysis is frequently interrupted by one or two ventures into unrelated aspects of the problem. This occurs particularly during the analysis of inhibitory relationships, where the subject must learn how to do something in order to prevent something else from occurring which prohibits solution. As mentioned earlier, this is the greatest source of difficulty in the problems. (Perhaps the sallies into unrelated areas which occur so often during this phase of the problem are to be considered as some sort of random attempt to avoid continued failure here. Often the unrelated actions are such as to elicit phenomena perfectly well known to the subject, and give somewhat the impression that after encountering frustration in the analysis of a difficult relation he consoles himself briefly by demonstrating his ability to control some aspects of the situation.)

Repetition of particular questions and elicitation of the same information in a number of different ways occurs to an extent which is surprising in a university population. Occasionally one observes repeated patterns of inquiry which are long and complicated but which appear over and over during the process. The over-all efficiency of the population, in terms of maximum utilization of the information content of events observed, is remarkably low. Even though subjects have decided to minimize questions rather than time (the instructions, the reader may recall, request that both be minimized), the use of inference *to avoid* asking a question is minimal. Rather, inference is tested *by* asking the question. Subjects give the impression of being unwilling to test an inference by applying the results to an attempt to achieve solution, as though they somehow attached more importance to such solution attempts than to other questions. Willingness to test an inference by using it in a synthetic attempt might be considered an index of willingness to take risk.

Separate analytic and synthetic phases can be clearly distinguished in the great majority of PSI performances. Generally, the clearer this distinction, the more effective is the performance. The shift to synthesis usually occurs prematurely; there is a tendency to try to make-do with insufficient information. Subjects will come back to analysis when forced to do so by their inability to produce a solution. The shift to synthesis occurs after surprisingly few questions in many cases.

Although performance is predominantly analytic, the shift point is located early enough on a relative scale to make clear that the necessity to reconcile simultaneously three constraints, which is the logical requirement imposed in the two problems used in this work, is exceedingly difficult for this population. The synthetic aspect of the task contributes almost as much to the difficulty as does the analytic component.

Frequently, subjects were asked to describe the way in which their manipulations of the PSI contributed to the achievement of the output. It was observed that when synthesis was carried to a successful conclusion without a complete analysis preceding it, the subject would have difficulty in repeating his solution (which was always required) and might have a solution which was not parsimonious, i.e., a solution which contained irrelevant manipulations.

In such cases, frequently the subject would attribute completely erroneous functions to some aspects of the solution combination. When inquiry was pursued further, the subject would buttress his rationale by citing the existence of relationships which he had observed repeatedly during the analysis. On a number of occasions, the evidence so adduced was diametrically opposed to reality. Thus one gains the impression of a perceptual factor in the PSI situation, which enables a subject to see one thing, conclude that he saw the opposite, and subsequently experience the repeated contradiction of his conclusion by observed phenomena without revising the way in which he has structured the situation.

B. Some Effects of Increasing Difficulty

As the problem becomes more difficult, by what is essentially the addition of a single feedback relationship, a number of changes occur in the characteristic process. The time required for solution to be achieved almost doubles and the number of questions asked becomes 1.6 times as great. The variety of these questions actually decreases, so that we have a tendency toward greater stereotypy of response. This increase in stereotypy is such that where an item of information will be elicited about 5.4 times in the average performance on the easier problem, it will be elicited about 8.4 times on the more difficult problem. Not only does redundancy increase, but something like "rigidity" increases also, as indicated by a tendency toward perseveration of useless activity on the more difficult problem. These changes occur in spite of the fact that subjects almost invariably take notes.

The complexity of the average question increases markedly, even though the relationships being analyzed are of the identical complexity in both problems. The average rate decreases appreciably, and the percentage of time spent with no manipulation of the PSI doubles. The number of changes of set goes up, and the amplitude of the effect of these changes increases.

Actual redundancy increases from 79% to 86%. The fact that an over-all increase of a factor of 0.6 in the total number of questions asked is accompanied by only 7% increase in redundancy is some indication of the extent to which analyzing the ordering of a process generates conclusions other than those obtained by more static methods of treating data. The absolute inferential lag increases by about 50%, but this lag represents about 50% of the total performance in both cases. In other words, about one-half of the performance is occupied by the elicitation of relationships which could be inferred on the basis of information already obtained, before the necessary and sufficient information (NASI) which enables a solution is acquired. Further superfluous questions may, of course, be asked after this point, and invariably are.

The absolute synthetic lag is negative in both problems. The average subject tends to shift to synthesis before he obtains the NASI (Explicit) but subsequently finds it necessary to revert to analysis to obtain this. This premature shift is greater on the more difficult problem. It is interesting that the shift to synthesis occurs in both problems at about the same point after the achievement of the NASI (Implicit), although of course this shift is relatively earlier on the more difficult problem. The similarity in the absolute locus of the shift appears almost as though shift occurs after a fixed and characteristic amount of analysis "whether the problem is ready or not."

The number of inversions from synthesis back to analysis doubles with the increase in difficulty, the performance becomes much less predominantly analytic, and the mixture of analysis and synthesis before the shift point is greater. All of these considerations support the conclusion that as the difficulty of the problem is increased, this group of subjects experienced relatively greater obstacles in that aspect of the task which required the synthesis of known relationships into an integrated product which met the definition of a solution.

Thus it appears that some aspects of the PSI performance are fairly constant and characteristic, while others are appreciably modified by increments in difficulty. The relatively invariant aspects of the behavior might reflect the intrinsic organization of the individual performance, and might give an insight into the extent to which factors related to personality influence the problem-solving attempt. The more labile aspects of performance may be consequences of reorganizations of process under the impact of greater stress.

IV. Applications of the PSI

A. Comparisons Between Groups

In order to clarify the factors which contribute to performance on the various Problem-Solving and Information Apparatus (PSI) variables, we have presented data collected from small groups of different composition with respect to educational level and area of specialized training or interest. These data have been used to estimate the dependence of our variables on these factors throughout the earlier sections of this paper. However, we may also use this material as a basis with which to evaluate the over-all contribution of education in general and of various kinds to PSI performance as a whole.[6]

A number of questions immediately come to mind when one considers the nature of the task involved in the PSI. Will education as such improve performance, and in what ways? Is training in the natural sciences more effective than training in other areas, and how is such increased effectiveness manifested? Can differences be demonstrated which are not dependent on such special skills of logical manipulation as are acquired in a university? What aspects of real-life performance are related to the behavior elicited by the PSI? Can one distinguish different modes of behavior which are concomitants of differences in personality structure?

The results which we have obtained clarify some of the questions asked above. We cannot claim to have provided definitive answers to many, nor did we expect to do so in as frankly exploratory a

study as this. Many of our comparisons are made with small samples which show consistent trends but do not reach statistical significance. Yet a clear picture begins to emerge from the data. Most of all, it becomes apparent that answers to these questions can be elicited using this technique, and areas which merit additional intensive inquiry become highlighted.

The over-all effects of education, as demonstrated by the results of a comparison between small groups of first- and eighth-year university students across our full set of variables, are to improve performance. Yet this improvement is so slight as to suggest that much more is involved than simply education per se. A comparison of Ph.D.-level individuals in the natural sciences with Ph.D.-level individuals in other areas of specialization shows that, on *each* index of performance, the natural-scientist group differs from the other group in a direction of greater effectiveness. This raises two possibilities: Either training in the natural sciences is the basis for superior performance of a group so trained, or perhaps individuals with a particular kind of habitual approach well-suited to this kind of problem are more likely to enter the natural sciences.

In an effort to ascertain whether the latter factor was operative, a group of college students was divided into two subgroups, one of which intends presently to enter the natural sciences for specialized study, and another which does not. Comparison of these two groups shows that in almost every respect, *the differences demonstrable between the two groups at the Ph.D.-level already exist between these two groups at the college level, before any specialized training.* It appears that an individual is likely to have a predictable approach and general behavior on the PSI based on his career interests. If one assumes that proficiency in the generic skills required in an area increases the probability that one will enter that area, it would appear that the activities involved in effective PSI performance are more related to the natural sciences than to other academic disciplines. The college-level group which intends to enter the natural sciences displays more effective behavior in a large number of variables than does the non-natural-scientist Ph.D.-level group.

If one compares groups with interests in the natural sciences at high and low educational levels, one can in effect partial out the contribution of differential interests to obtain an approximation of the effects of natural-science education on PSI performance, and similarly with groups in other areas of endeavor. The results of such an evaluation of the impact of different disciplines on PSI performance show that, while there is a definite quantitative increase in effectiveness after training, that increase is not as great as one would hope, does not span all variables measured, and is not accompanied by a qualitative change in the problem-solving process. That is, one concludes that for the present population, which is by no means representative, and for the PSI, which is a particular kind of problem, *advanced education serves to improve performance along pre-existing lines rather than to establish new and more effective problem-solving techniques,* and the amount of this improvement is not great.

The data raise the question of the developmental factors which contribute to the establishment of the differences which we observe before specialized training. A small number of experiments, which will not be reported in detail, carried out with extremely young (6- to 11-year-old) children, has shown that such technically naive subjects not only can solve PSI problems, but that the quality of some of their performances was better than that observed with some subjects *at the Ph.D. level.* It would be of great interest to determine whether two groups emerge in a population of children after a given level of experience or education is reached, whether such factors as home environment can be related to differential PSI performance, or whether perhaps two groups are found in the earliest measurements which can be made.

B. Relation Between PSI, Personality, and College Performance

The PSI would appear to require exercise of skills in analysis and synthesis. The results obtained from cross-disciplinary comparisons suggest that this is indeed the case, if one assumes that the natural sciences require skill in analysis and synthesis. In an attempt to obtain a more direct demonstration of the relation between facility with the PSI and facility in the performance of real-life tasks which require analysis and synthesis, we analyzed the performance of 16 college students on all comprehensive examinations taken during their residence at the University of Chicago. These examinations have subtests which measure the ability of the subject to elucidate and apply principles in the area of the examination. Examining the relationship between certain variables of PSI performance and scores on all such examination subtests in the subject's college career to date, we find a clear correspondence between effectiveness on the

PSI and the ability to do well on examination sections which require analysis and synthesis. Based on a trichotomy of high, medium, or low performance on the criterion of actual redundancy on the more difficult problem, a better ordering was obtained between relative college grades on the one hand, and cumulative analytic and synthetic subtest scores on the other, than by using either ACE scores or college grades alone.

Finally, preliminary examination of the relationship between differential PSI performance and some aspects of personality, carried out on the same sample of 16 cases by Mr. Sidney Blatt of the Psychology Department of the University of Chicago, shows a clear relationship between these factors.[6]

V. Summary and Conclusions

A technique has been developed and herein described and discussed which permits the detailed observation of the problem-solving process rather than its product. The technique involves the use of an apparatus which permits the presentation of many different abstract problems of quantifiable information content and structure. This apparatus is called the Problem-Solving and Information Apparatus (PSI).

A technique of constructing and scoring problems using the PSI has been described and discussed which enables objective measurement of the phases of acquisition, organization, and application of information by an individual during the process of achieving a solution to abstract problems. These aspects of the problem-solving process are related to the solution in a causal fashion, so that the solution is seen to be the outcome of a unified and cohesive sequential process.

Using 59 university students and staff members at different educational levels and with different areas of specialization, the dependence of variables in the areas of work habits, information acquisition and handling, and consistency and appropriateness of approach to the problem, on such factors as education, specialized technical training, and career interest has been indicated.

To the extent that real-life facility in the solution of problems of the generic sort represented by the task which is posed by the PSI may be inferred from observation of PSI performance, a number of conclusions are warranted by the data. First, marked differences exist between groups of individuals engaged in the advanced study of certain disciplines. Second, these differences exist qualitatively between groups intending to enter these disciplines, before specialized study has occurred. Third, the effect of education varies according to the discipline studied, with performance in the various areas of the problem-solving process being affected differentially.

The relationship between PSI performance and ability to elucidate and apply principles in college courses has been demonstrated to be appreciable. It would appear that difficulty with such manipulations constitutes a major and severe constraint on the performance of all groups studied.

The data indicate that training or experience in certain activities, whether acquired early in life or during the course of academic studies, bring about habituation of an individual to certain kinds of conceptual and organizational processes which are consistently displayed in repeated PSI performance. These processes may be more or less appropriate to the requirements imposed by this sort of problem. In addition to such habitual processes, some aspects of personality appear to be reflected in the problem-solving process, such as, for example, self-confidence, anxiety, and compulsiveness. Personality factors as well as cognitive factors contribute to the PSI performance.

Further investigations of the problem-solving process using this technique are necessary before definitive analyses of the factors involved in performance can be attempted. Many potential applications must await the completion of such research. The technique may be useful for the development and evaluation of teaching and remedial techniques in certain fields, by evaluation of the performance of individuals before and after training in particular skills. By appropriate construction of parallel forms both containing and not containing specialized information, it might be possible to evaluate performance in certain professional and technical areas in a diagnostic way, so as to ascertain the extent to which difficulty in the manipulation of abstractions rather than deficiency in knowledge was the major constraint on the technical competence of an individual. The relationship of performance on this task to performance on standard psychological assessment devices must of course be determined, but perhaps of equal interest would be elucidation of the contribution of the kind of behavior here observed to abilities and skills used in less artificial situations. The method might be useful in evaluating the effectiveness of chemotherapy or other procedures on the ability of disturbed individuals to organize cognitive processes.

In view of the clear differences which have been demonstrated between different groups entering college, and the relation of these differences to subsequent choice of career, it would be particularly interesting to gain an understanding of the factors which contribute to the establishment of these differences and the acquisition of these skills, the age level at which they are acquired, and the kinds of experience which are most conducive to their acquisition and effective retention. The results of such developmental studies might enable more appropriate preparation for those areas of intellectual endeavor where this sort of skill is of importance, by facilitating the construction of generalized methods for increasing effectiveness on tasks of this generic type. Present academic training does not appear to change parameters of effectiveness qualitatively to any appreciable extent.

Finally, the possibility of devising simpler tasks for laboratory animals, which permit the same sort of sequential observation and analysis of cognitive behavior, should be investigated. Present behavioral indices tend to be product measures, and possible data of value may be obscured by the failure to construct indices which give more insight into process itself.

BIBLIOGRAPHY

1. Bloom, B. S., & Broder, L. J. Problem-solving process of college students. *Suppl. educ. Monogr.*, No. 73, July 1950.
2. Duncker, K. On problem-solving. (Trans. by Lynne S. Lees.) *Psychol. Monogr.*, 1945, **58**, No. 5 (Whole No. 270).
3. Glaser, R., Damrin, D. E., & Gardner, F. M. The tab item: A technique for the measurement of proficiency in diagnostic problem solving tasks. *Educ. psychol. Measmt*, 1954, **14**, 283.
4. Grings, W. W. *A methodological study of electronic trouble shooting skill.* Rep. No. 9, Contract NONR 282-02, USC, LA, 1953.
5. Harlow, H. F. *Symposium on psychology of learning basic to military training problems.* Rep. No. HR-HTD 201/1, Panel on Training and Training Devices, Comm. on Human Resources, Res. & Dev. Bd, May 1953, pp. 165–78.
6. John, E. R., & Rimoldi, H. J. A. Sequential observation of complex reasoning. *Amer. Psychologist*, 1955, **10**, 470. (Abstract)
7. Rimoldi, H. J. A. A technique for the study of problem-solving. *Educ. psychol. Measmt*, 1955, **15**, 450.

(Accepted for publication March 14, 1957)

GEORGE BANTA COMPANY, INC., MEN[illegible]